Gouvernement du Canada
Ministère des Communications

Government of Canada
Department of Communications

Le Centre canadien de recherche sur l'informatisation du travail
Canadian Workplace Automation Research Centre

CA1
DTR290-
90
R26Z
v.1

41.D

Canada

CA1
DTR290-
90
R26Z
v.1

472120

41.D

Communications Canada
Centre canadien de recherche sur l'informatisation du travail (CCRIT)
Direction des Systèmes Intégrés

92-05161/

TP 10575F

REMPLACEMENT D'UN PSEUDO-PILOTE PAR DES TECHNOLOGIES DE RECONNAISSANCE ET DE GÉNÉRATION DE LA PAROLE

VOLUME I

pour le
Centre de développement des transports
Groupe des politiques et de la coordination
Transports Canada

Juillet 1990

472120

DDN 456177

TP 10575F

Remplacement d'un pseudo-pilote par des technologies de reconnaissance et de génération de la parole

Volume I

par

Jean-Yves Fiset

École Polytechnique de Montréal et
Centre canadien de recherche sur
l'informatisation du travail

Juillet 1990

AVERTISSEMENT : Les opinions et les vues exprimées dans ce rapport sont celles de l'auteur et ne reflètent pas nécessairement celles du Centre de développement des transports, Transports Canada.

Also available in English under the title "Replacing a pseudo-pilot by speech generation and recognition technology", TP 10575E.

Transports Canada Transport Canada

FORMULE DE DONNÉES POUR PUBLICATION

1. N° de la publication de Transports Canada	2. N° de l'étude	3. N° de catalogue du destinataire
TP 10575F	7366	

4. Titre et sous-titre	5. Date de la publication
Remplacement d'un pseudo-pilote par des technologies de reconnaissance et de génération de la parole Volume I	Juillet 1990
	6. N° du document de l'organisme exécutant

7. Auteur(s)	8. N° de dossier — Transports Canada
Jean-Yves Fiset	ZCD 1455-201-2

9. Nom et adresse de l'organisme exécutant	10. N° de dossier — ASC
Centre canadien de recherche sur l'informatisation du travail 1575, boul. Chomedey Laval (Québec) H7V 2X2	11. N° de contrat — ASC ou Transports Canada

12. Nom et adresse de l'organisme parrain	13. Genre de publication et période visée
Centre de développement des transports 200 ouest, boul. René-Lévesque Tour ouest, suite 601 Montréal (Québec) H2Z 1X4	Rapport final
	14. Agent de projet James D. Reid

15. Remarques additionnelles (Programmes de financement, titres de publications connexes, etc.)
Volume II : Annexes

16. Résumé

But : ce rapport vise à déterminer comment on peut remplacer un pseudo-pilote humain par un dispositif utilisant des techniques de reconnaissance et de génération de la parole.

Méthode : on examine d'abord l'état actuel de la recherche appliquée en Amérique du Nord et en Europe. On élabore ensuite un modèle sur mesure du langage des élèves contrôleurs. On examine alors l'utilisation des technologies de reconnaissance et de génération de la parole.

Résultats : on constate qu'il y a un certain nombre de systèmes en cours de développement, même si aucun n'est encore en exploitation. On propose la représentation par schémas comme modèle de langage pour un pseudo-pilote automatisé. On formule des recommandations sur l'exploitation des technologies de reconnaissance et de génération de la parole.

Conclusion : il semble techniquement faisable de réaliser un pseudo-pilote automatisé. On doit cependant choisir si on veut un système passif (i.e. un simulateur commandé par la voix) ou un système actif (i.e. qui conseillera l'élève contrôleur). On doit aussi choisir si on veut un système réalisé entièrement au Canada ou si on désire établir une coopération avec un organisme qui développe actuellement des pseudo-pilotes automatisés.

17. Mots clés	18. Diffusion
Contrôle aérien, pseudo-pilote, langage, technologies vocales, intelligence artificielle.	Le Centre de développement des transports dispose d'un nombre limité de copies.

19. Classification de sécurité (de cette publication)	20. Classification de sécurité (de cette page)	21. Déclassification (date)	22. Nombre de pages	23. Prix
Non classifiée	Non classifiée	—	xvii+110 +annexes	—

CDT/TDC 79-005
Rev 88

Canada

Transport Canada / Transports Canada

PUBLICATION DATA FORM

1. Transport Canada Publication No.	2. Project No.	3. Recipient's Catalogue No.
TP 10575F	7366	

4. Title and Subtitle	5. Publication Date
Remplacement d'un pseudo-pilote par des technologies de reconnaissance et de génération de la parole Volume I	July 1990
	6. Performing Organization Document No.
7. Author(s)	8. Transport Canada File No.
Jean-Yves Fiset	ZCD 1455-201-2
9. Performing Organization Name and Address Canadian Workplace Automation Research Centre 1575 Chomedey Blvd. Laval, Quebec H7V 2X2	10. DSS File No.
	11. DSS or Transport Canada Contract No.
12. Sponsoring Agency Name and Address Transportation Development Centre 200 René-Lévesque Blvd. West West Tower, suite 601 Montreal, Quebec H2Z 1X4	13. Type of Publication and Period Covered Final
	14. Project Officer James D. Reid

15. Supplementary Notes (Funding programs, titles of related publications, etc.)
Volume II : Appendices

16. Abstract

Objective : The objective of this report is to determine how to replace a human pseudo-pilot with a device using speech recognition and generation technology.

Method : First, the state of the applied research in North America and Europe is surveyed. Then, a custom language model for air traffic control is developed. Finally, the use of speech recognition and generation technology for this application is reviewed.

Results : Although several systems are in the development stage, none is in current use. The use of a frame-based language model is recommended in this application. Recommendations are made regarding the use of speech recognition and generation technology.

Conclusion : The development of a voice driven pseudo-pilot seems technically feasible. However, a choice must be made between a passive system (e.g. voice-driven simulator) and an active system (e.g. one that will coach the trainee). A choice must also be made between an entirely Canadian system and a system developed in cooperation with a partner already active in this field.

17. Key Words	18. Distribution Statement
Air traffic control, pseudo-pilot, language, speech technology, artificial intelligence	Limited number of copies available from the Transportation Development Centre on request.

19. Security Classification (of this publication)	20. Security Classification (of this page)	21. Declassification (date)	22. No. of Pages	23. Price
Unclassified	Unclassified	—	xvii+110 +annexes	—

REMERCIEMENTS

L'auteur désire remercier les personnes suivantes pour l'aide qu'elles lui ont apporté lors des différentes phases de ce projet :

M. James D. Reid, du Centre de développement des transports, Groupe des politiques et de la coordination, Transports Canada, qui nous a fourni l'opportunité de travailler sur ce projet.

M. Raymond Descout, du Centre canadien de recherche sur l'informatisation du travail, de Communications Canada, qui nous a fourni un environnement de travail parmi les meilleurs et qui nous a fait profiter de son expérience dans l'exploitation des technologies vocales.

M. Jean-Marc Robert, de l'École Polytechnique de Montréal, qui nous a encouragé sur la voie de la recherche et qui a assuré l'encadrement académique de ce projet.

Mme Françoise Néel du Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (LIMSI) d'Orsay (France), qui a aimablement organisé une visite de son laboratoire.

M. Michael J. A. Asbury du Civil Aviation Authority (Grande-Bretagne), qui a prêté son concours pour organiser une visite de plusieurs centres de recherche en Angleterre.

Il n'est pas possible de nommer les nombreuses autres personnes qui ont contribué de près ou de loin à ce projet, mais dans tous les cas où elle s'est manifestée, cette aide a été sincèrement appréciée.

SOMMAIRE

Ce rapport présente les résultats d'une étude sur la faisabilité de remplacer un pseudo-pilote utilisé pour la formation des contrôleurs de la circulation aérienne par un dispositif utilisant des techniques de reconnaissance et de génération de la parole.

La première étape du projet consiste en un tour d'horizon de la recherche appliquée en Amérique du Nord et en Europe. Au Canada, la société ATS a développé un pseudo-pilote qui peut accepter des commandes vocales dans un environnement de contrôle radar. Bien qu'aucun système de ce genre n'ait encore été mis en service, ATS a réalisé un démonstrateur. Aux États-Unis, la FAA prévoit évaluer deux systèmes :

- Le système TOTS (Tower Operator Training System) de la société Logicon : il s'agit d'un système développé initialement pour l'Aéronavale américaine. On prévoit l'expédier au centre de formation de la Federal Aviation Administration en 1990; la mise en marche est prévue en 1991. Le système sert aux exercices de contrôle en vol à vue.

- Le système ATCoach, de la société UFA : il s'agit d'un système utilisant des techniques d'intelligence artificielle auquel on peut adjoindre une entrée vocale. Le système sert aux exercices de contrôle radar.

Pour l'Europe, il semble que les deux intervenants principaux soient la France et la Grande-Bretagne. En France, une collaboration entre le CENA (Centre d'études de la navigation aérienne), le LIMSI (Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur, rattaché au Centre national de la recherche scientifique) et la société Vecsys a conduit à la réalisation d'un prototype de pseudo-pilote qui semble très performant au point de vue de la reconnaissance de la parole. En Grande-Bretagne, on prévoit installer au Air Traffic Control Evaluation Unit de Bournemouth un pseudo-pilote hybride utilisant à la fois la reconnaissance de la parole et une entrée tactile. Ce système ne s'adresse pas à la formation d'élèves contrôleurs débutants; des contrôleurs d'expérience l'utiliseront pour des exercices plus avancés.

Pour la seconde étape du projet, on développe un modèle sur mesure du langage utilisé dans les exercices de simulation dans le contexte canadien. Ce modèle s'inspire de celui utilisé par le LIMSI et a été élaboré à partir de transcriptions d'exercices enregistrés au Centre de formation de Transports Canada à Cornwall. Dans ce modèle, on tente d'associer chacun des messages des élèves contrôleurs à une catégorie donnée. Cette catégorie constitue un schéma, ce qui est une façon de représenter des connaissances en intelligence artificielle. Ce genre de modèle permet de pallier les limitations des systèmes actuels de reconnaissance de la parole et facilite la mise au point d'un pseudo-pilote automatisé intelligent. On recommande d'utiliser ce genre de modèle dans un pseudo-pilote automatisé, en y incorporant des processus appropriés de gestion des erreurs de dialogue.

La dernière étape de cette étude consiste à examiner l'utilisation des technologies de reconnaissance et de génération de la parole, en tenant compte:

- Des caractéristiques de la tâche de l'élève contrôleur.
- Des caractéristiques de la technologie disponible.
- Des aspects ergonomiques.

Ainsi, on recommande l'emploi de voix numérisée pour la génération des réponses du pseudo-pilote automatisé ainsi que l'utilisation d'un reconnaisseur de parole opérant en mots connectés et monolocuteur. On recommande aussi d'incorporer l'obtention des échantillons de voix nécessaires pour le reconnaisseur de parole à l'apprentissage de la phraséologie.

En conclusion, au-delà de la faisabilité technique d'un pseudo-pilote automatisé, on soutient qu'il est important pour la suite du projet de résoudre les questions suivantes:

- Qu'est-ce qu'on veut obtenir d'un pseudo-pilote automatisé? Il faut déterminer si le pseudo-pilote sera passif dans la formation des élèves contrôleurs (i.e. il exécute les directives valides), ou s'il sera actif (i.e. il conseillera l'élève contrôleur en plus d'exécuter les commandes valides).

- Tient-on à réaliser un système qui soit entièrement canadien? Selon la réponse à la première question (pseudo-pilote passif ou actif), il peut être avantageux d'établir une forme de coopération avec des organismes qui ont acquis une expérience considérable dans le domaine.

TABLE DES MATIÈRES

1. INTRODUCTION 1

2. ÉTAT DE LA RECHERCHE APPLIQUÉE ACTUELLE 3
   - 2.1 Canada 3
   - 2.2 États-Unis 4
   - 2.3 Europe 10
   - 2.4 Synthèse de la recherche appliquée actuelle 14

3. CHOIX D'UNE SITUATION DE CONTRÔLE AÉRIEN POUR L'ÉLABORATION D'UN MODÈLE DE LANGAGE 16
   - 3.1 Discussion sur la phraséologie utilisée dans les différentes situations de contrôle aérien 16
   - 3.2 Description de la situation considérée dans cette étude 17
   - 3.3 Description du poste de l'élève contrôleur 20

4. CHOIX DU MODÈLE DE REPRÉSENTATION DU LANGAGE CONTRÔLEUR-PILOTE 24
   - 4.1 Analyse de modèles existants 24
   - 4.2 Choix du modèle de représentation du langage 29

5. DÉMARCHE DE LA MODÉLISATION DU LANGAGE OPÉRATIF 31
   - 5.1 Dialogue cible 32
   - 5.2 Description de la démarche de développement du modèle 32

6. PREMIERE ÉTAPE DU DÉVELOPPEMENT DU MODÈLE POUR DES EXERCICES SIMPLES 37
   - 6.1 Examen de la longueur des transmissions 37
   - 6.2 Obtention du vocabulaire 38
   - 6.3 Élaboration des schémas préliminaires 42
   - 6.4 Validation de l'utilisation des schémas dans le cas du même exercice exécuté par des élèves contrôleurs différents 44
   - 6.5 Représentation d'un appareil 46
   - 6.6 Relation entre le dialogue et les états d'un appareil 47

7. SECONDE ÉTAPE DU DÉVELOPPEMENT DU MODÈLE POUR LES EXERCICES SIMPLES 50
7.1 Règles d'instanciation des schémas 50
7.2 Gestion des erreurs de dialogue 55
7.3 Évocation de schémas successifs 63
7.4 Modèle résultant 67

8. EXTENSION DU MODÈLE RÉSULTANT 71
8.1 Adaptabilité du modèle 71
8.2 Relâchement des contraintes sur la concaténation des messages et le nombre de destinataires d'une transmission 77
8.3 Remarques sur la modélisation du langage 81

9. EXIGENCES ET POSSIBILITÉS POUR LA RECONNAISSANCE DE LA PAROLE 82
9.1 Description des caractéristiques requises pour le reconnaisseur de parole 82
9.2 Apprentissage avec le reconnaisseur de parole 84
9.3 Effet de l'interaction avec une machine sur le comportement langagier 86

10. EXIGENCES ET POSSIBILITÉS POUR LA GÉNÉRATION DE LA PAROLE 87
10.1 Particularités pertinentes de la tâche 87
10.2 Description et caractéristiques des méthodes de génération 88
10.3 Aspects ergonomiques: perception de la parole générée 93
10.4 Effets de l'utilisation de parole générée artificiellement sur le comportement langagier 98
10.5 Choix d'une voix synthétique pour le pseudo-pilote automatisé 100

11. CONCLUSION 103

ANNEXES

Annexe A: Extrait du plan de cours de l'Institut de formation de Transports Canada, Cornwall, Ontario

Annexe B: Transcriptions d'exercices simples

Annexe C: Vocabulaire des élèves contrôleurs pour les exercices simples

Annexe D: Élaboration des schémas individuels pour les exercices simples

Annexe E: Résumé des schémas et des sous-schémas pour les exercices simples

Annexe F: Attribution des cotes

Annexe G: Analyse des processus de gestion des erreurs

Annexe H: Transcriptions d'exercices complexes

Annexe I: Vocabulaire des élèves contrôleurs pour les exercices complexes

Annexe J: Ébauche de schémas additionnels pour les exercices complexes

Annexe K: Analyse de transmissions concaténées

Annexe L: Vocabulaire des pseudo-pilotes

LISTE DES FIGURES

Figure 1	Schéma-bloc du système ATCoach	7
Figure 2	Schéma-bloc du pseudo-pilote développé au LIMSI	11
Figure 3	Zone de contrôle régionale de Midtown	18
Figure 4	Poste de travail de l'élève contrôleur	21
Figure 5	Représentation du langage sous forme d'arborescence	25
Figure 6	Représentation du langage sous forme de schémas	27
Figure 7	Représentation schématique de l'analyse sommaire du dialogue	34
Figure 8	Représentation schématique de l'amélioration du modèle de dialogue	36
Figure 9	Distribution de la longueur des phrases pour les exercices simples	38
Figure 10	Effet du dialogue sur l'évolution d'un appareil dans l'espace d'états	48
Figure 11	Squelette du système de gestion du dialogue	68
Figure 12	Exemple de schéma amélioré	69
Figure 13	Distribution de la longueur des phrases pour les exercices complexes	72
Figure 14	Comparaison de la distribution de la longueur des phrases pour les exercices simples et les exercices complexes	73
Figure 15	Étapes pour la concaténation de mots ou de segments de phrases	89

Figure 16 Éléments principaux d'un système de synthèse à partir du texte 92

LISTE DES TABLEAUX

Tableau I	Performance du reconnaisseur de parole SSI pour le système ATCoach	9
Tableau II	Résultats sur l'évaluation de la taille du vocabulaire pour les exercices simples	39
Tableau III	Évaluation du nombre d'éléments standards	41
Tableau IV	Taille du vocabulaire opératif pour les exercices simples	42
Tableau V	Exemple d'élaboration d'un schéma	43
Tableau VI	Représentation du schéma HDR	44
Tableau VII	Proportion des messages décrits par les divers schémas	45
Tableau VIII	Représentation des divers états d'un appareil	46
Tableau IX	Relations entre la représentation d'un appareil et les schémas	52
Tableau X	Limites pour l'instanciation de divers sous-schémas	54
Tableau XI	Classification des erreurs de dialogue	56
Tableau XII	Méthodes de gestion des autocorrections	58
Tableau XIII	Méthode de gestion des demandes de répétition originant de l'élève contrôleur	59
Tableau XIV	Méthode de gestion des demandes de répétition originant du pseudo-pilote	60
Tableau XV	Mécanismes d'identification et de résolution des ambiguités	61

Tableau XVI	Évocation de schémas successifs	65
Tableau XVII	Taille du vocabulaire opératif net (incomplet) pour les exercices de terminal	74
Tableau XVIII	Taille du vocabulaire opératif pour le contrôle de terminal	74
Tableau XIX	Vérification de l'utilité des schémas du modèle simple pour le contrôle de terminal	76
Tableau XX	Analyse des transmissions contenant plusieurs messages pour le même destinataire	78
Tableau XXI	Analyse des transmissions contenant des messages différents pour plus d'un destinataire	80
Tableau XXII	Vocabulaire des pseudo-pilotes pour toutes les transcriptions	88
Tableau XXIII	Sommaire de résultats sur des tests de netteté, pour des systèmes de synthèse à partir du texte	95

1. INTRODUCTION

Cette étude, réalisée dans le cadre d'un contrat de recherche octroyé par Transports Canada, vise à établir comment on peut remplacer un pseudo-pilote humain par un dispositif automatisé utilisant la reconnaissance et la génération de la parole. Un pseudo-pilote est un individu qui opère un simulateur suivant les consignes orales que lui donne un élève contrôleur de trafic aérien. Dans un centre de formation comme celui de Transports Canada à Cornwall, il y a plusieurs pseudo-pilotes qui contribuent à la formation des élèves contrôleurs. Plusieurs raisons peuvent être invoquées pour motiver la mise au point d'un pseudo-pilote automatisé :

- Les coûts associés (formation et salaires) à l'utilisation de pseudo-pilotes humains.

- Le manque de disponibilité de pseudo-pilotes qualifiés, soit en général, soit à demande par l'élève contrôleur quand ce dernier ressent le besoin de pratiquer une procédure particulière.

- La possibilité d'améliorer la standardisation de la formation des élèves contrôleurs.

Selon les termes de l'entente avec Transports Canada, cette étude examine successivement les aspects suivant :

- L'état actuel de la recherche appliquée dans le domaine des pseudo-pilotes automatisés. Pour des fins de commodité, on n'a considéré dans cette étude que la recherche appliquée effectuée en Amérique du Nord et en Europe.

- L'élaboration d'un modèle "sur mesure" du langage opératif de l'élève contrôleur, de façon à maximiser la performance d'un système de reconnaissance de la parole. Cette étape implique de choisir la situation de contrôle aérien pour laquelle on veut modéliser le langage. Il est important de noter que l'on ne vise pas à développer ici un système ou une stratégie de reconnaissance de la parole; on propose plutôt de

créer une représentation du langage opératif utilisé dans la réalité pour certains exercices qui soit exploitable par un système de reconnaissance de la parole.

- Les technologies de reconnaissance et de génération de parole. On tient compte des exigences de la tâche, des considérations ergonomiques et des caractéristiques des technologies qui sont disponibles.

Même si cette étude propose des réponses, on espère qu'elle suscitera également des questions de fond sur les aspects suivants :

- Le rôle qu'on veut faire jouer à un pseudo-pilote automatisé dans la formation des élèves contrôleurs.

- La pertinence de développer un pseudo-pilote automatisé qui soit entièrement canadien.

2. ÉTAT DE LA RECHERCHE APPLIQUÉE ACTUELLE

Ce chapitre fait le point sur l'état actuel de la recherche appliquée en Amérique du Nord et en Europe en ce qui concerne la réalisation de pseudo-pilotes automatisés (incluant certains produits commerciaux connexes). Pour y arriver, on examine un certain nombre de pseudo-pilotes automatisés qui ont été développés soit par des sociétés privées, soit par des organismes gouvernementaux, soit par une combinaison des deux. On s'intéresse spécifiquement à la modélisation du langage des élèves contrôleurs, à l'ergonomie du dialogue avec le pseudo-pilote et aux prototypes existants. On ne tente pas d'évaluer ici les systèmes de reconnaissance de la parole qui sont utilisés. D'une part, l'évaluation des reconnaisseurs de parole n'est pas bien maîtrisée aujourd'hui. D'autre part, le reconnaisseur de parole ne constitue qu'un des éléments d'un pseudo-pilote automatisé : on doit donc évaluer la performance du système dans son ensemble (incluant par exemple le modèle du langage), et non seulement la performance d'un seul élément.

2.1 Canada

On n'a retrouvé qu'une organisation qui mette au point des systèmes de simulation pour les élèves contrôleurs et utilisant une entrée vocale : ATS Inc.[1]. Cette société avait déjà été impliquée dans le développement d'un système de simulation pour les élèves contrôleurs (1)[2]. Plus récemment, ATS a développé un système pouvant utiliser une entrée vocale. Selon nos informations (2), le système est utilisé pour entraîner des élèves contrôleurs pour une situation de contrôle radar. Au point de vue matériel, on utilise un PC 80286. Pour la reconnaissance de la parole, le système a les caractéristiques suivantes :

---

[1] ATS Aerospace Inc., 3163 rue Harvey, St-Hubert (Québec) J3Y 3T7, (514) 676-6259

[2] Les numéros entre parenthèses renvoient à la liste des ouvrages cités en fin de rapport.

- Reconnaisseur de parole Verbex, en mots connectés.

- Le reconnaisseur utilise une syntaxe à états finis, i.e. une structure arborescente décrivant les "trajets" possibles pour une phrase syntaxiquement valide. Le vocabulaire est de l'ordre de 300 mots.

Pour la génération de parole, le système peut utiliser soit une voix synthétique, soit une voix numérisée. Selon la même source, ATS n'a pas encore installé un système de ce genre (avec entrée vocale). Il existe toutefois un système de démonstration avec cette option.

2.2 États-Unis

Diverses organisations gouvernementales (civiles et militaires) sont intéressées au développement d'un pseudo-pilote automatique ou de systèmes connexes. Pour cette étude, nous nous intéressons surtout aux recherches dans le domaine de l'aviation civile, sauf dans les cas où il est utile d'examiner des systèmes provenant de la recherche militaire.

La FAA (Federal Aviation Administration) est l'organisme civil qui régit le transport aérien aux Etats-Unis. La FAA compte aussi divers organismes qui s'intéressent à la formation des élèves contrôleurs et à l'utilisation de technologies vocales pour la mise au point de pseudo-pilotes automatiques.

Au sein de la FAA, on reconnaît généralement (3) qu'il y a un besoin pour améliorer la quantité de temps de simulation (et donc de pseudo-pilotes) mis à la disposition des élèves contrôleurs pour faire face aux besoins anticipés. On fait toutefois face à des contraintes de budget et de disponibilité de personnel. L'alternative consistant à utiliser des pseudo-pilotes automatisés devient alors attrayante. A la FAA, on croit possible d'utiliser un pseudo-pilote automatisé à diverses étapes de la formation d'un contrôleur (4) : au niveau élémentaire (à l'Académie de la FAA à Oklahoma City), à un niveau plus avancé (dans des centres de formation régionaux) et même lors de la formation continue des contrôleurs expérimentés. Il existe toutefois une restriction dans le cas de la formation dans les centres régionaux : il y a actuellement au sein de la FAA une réglementation qui force à utiliser une copie du

logiciel d'opération réel pour y effectuer la formation des élèves contrôleurs. Un poste autonome de pseudo-pilote automatisé ne peut donc pas être exploité dans le cadre de la réglementation actuelle. En dépit de cette restriction pour les centres régionaux, on envisage d'évaluer différents types de pseudo-pilotes automatisés. Il y a actuellement deux filières distinctes pour la réalisation de pseudo-pilotes commandés par la voix et qui intéressent la FAA :

. Un système développé initialement pour le compte de l'Aéronavale américaine par la firme Logicon[1].

. Un système développé par la firme UFA[2] qui est un système-expert servant à entraîner des élèves contrôleurs, auquel on adjoint des technologies vocales.

On décrit ces deux systèmes de façon plus détaillée dans les paragraphes suivants.

2.2.1 Systèmes développés par Logicon

Logicon a été impliqué depuis plusieurs années avec l'Aéronavale américaine pour le développement de divers systèmes de pseudo-pilotes utilisant la reconnaissance de la parole (5), (6). Vers 1985, l'Aéronavale américaine lançait un nouvel appel d'offres avec les objectifs suivants (projet ASATS/SATS) :

. Réaliser un pseudo-pilote automatisé.

. Améliorer la standardisation de la formation du contrôle aérien.

La firme Logicon obtint ce contrat. Le système est un pseudo-pilote automatisé utilisé en contrôle radar. Peu après (en 1987), un second appel d'offre de l'US Navy lança le projet TOTS

---

[1] Logicon Inc., 4010 Sorrento Valley Blvd, San Diego, CA, 92121, (619) 455 1330

[2] UFA Inc., 335 Boylston St., Newton, MA, 02159, (617) 964 5172

(Tower Operator Training System); Logicon obtint aussi ce contrat. Par la suite, la FAA acquit un système TOTS pour fins d'évaluation. Logicon prévoit expédier ce système à la FAA vers décembre 1990. Le système devrait être opérationnel en 1991. Au point de vue matériel, le système est basé sur des stations SUN. On peut trouver une description plus complète dans (7).

Même si les deux systèmes diffèrent (contrôle de tour vs contrôle radar), on note qu'ils partagent certaines caractéristiques :

- Le système de reconnaissance de parole qui est utilisé est le VRS 1280 de ITT. Il s'agit d'un reconnaisseur de parole continue monolocuteur avec un vocabulaire possible de 500 mots (on peut répéter des mots dans la syntaxe jusqu'à concurrence d'un total de 2000 mots). On peut utiliser la même carte pour générer une voix artificielle (enregistrée préalablement avec un algorithme de compression CVSD).

- La reconnaissance utilise une syntaxe à états finis.

Selon le fournisseur du reconnaisseur de parole utilisé dans le système, ces systèmes partagent aussi un certain nombre de difficultés (8). On relève ici celles qui sont pertinentes à une application civile :

- Il peut y avoir plusieurs directives contenues à l'intérieur d'une transmission pour un même appareil.

- Il peut y avoir des messages différents émis à l'intérieur d'une transmission, et destinés à plus d'un appareil.

- Certaines des transmissions sont assez longues (de 30 à 40 mots).

- Il faut tenter de gérer les corrections que les élèves contrôleurs introduisent dans les messages.

Les exigences de l'Aéronavale américaine pour la reconnaissance de parole (9) sont de 98 % au niveau des mots, et de 94 % pour les phrases. Selon des évaluations internes chez Logicon, le système atteint actuellement 96.4 % de reconnaissance, au niveau des phrases. Pour atteindre ce niveau de performance, on a recours à des traitements internes dans le pseudo-pilote, pour

corriger certaines erreurs de reconnaissance.

2.2.2 Système développé par UFA

La firme UFA a développé un système (ATCoach) qui peut être utilisé pour la formation des contrôleurs aériens (10). Le développement de ce système a débuté vers 1984 à la suite d'un appel d'offre de la FAA. A la base, il s'agit d'un système de formation pour le contrôle aérien qui fait appel à des techniques de simulation et d'intelligence artificielle. L'élève contrôleur peut dialoguer avec le système soit manuellement, soit oralement. La figure qui suit montre le système (adapté de (11)) :

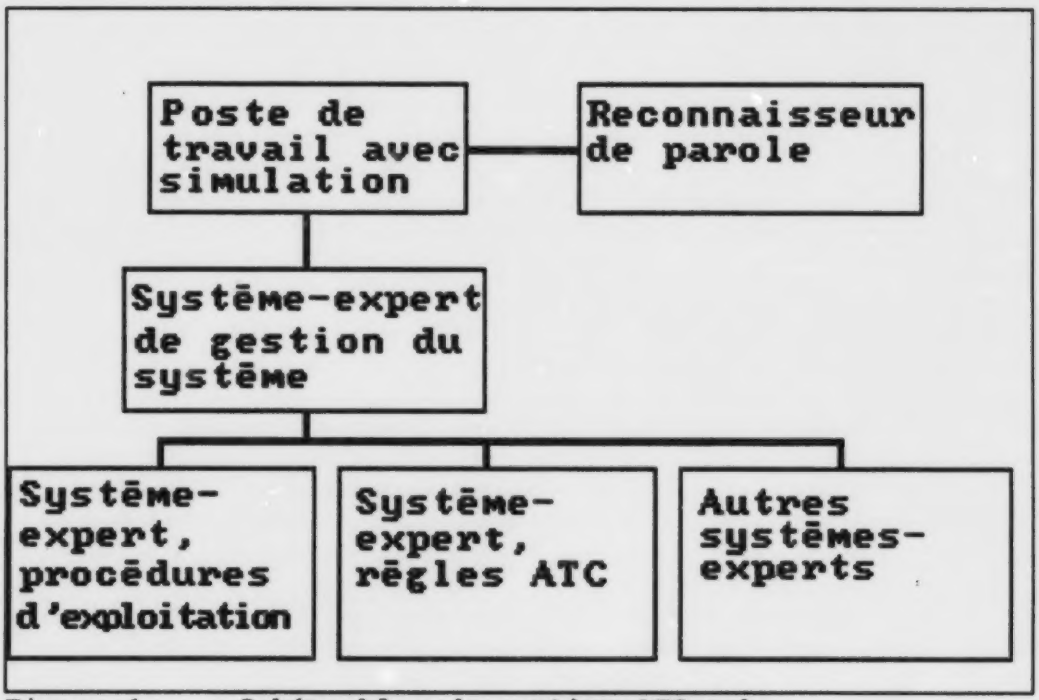


Figure 1 **Schéma-bloc du système ATCoach**

Lors d'un exercice, le système reçoit des directives de l'élève contrôleur et les évalue avant de les exécuter. Le cas échéant, un système-expert avise l'élève contrôleur si sa directive enfreint les procédures établies.

Il est possible d'avoir une station indépendante, ou plusieurs stations reliées en réseau pour permettre les transferts entre secteurs de contrôle. La plateforme utilisée pour la version la plus simple de ATCoach est un PC utilisant un processeur 80386, cadencé à 33 MHZ. Pour des versions plus élaborées, on peut faire appel à des postes de travail Apollo (12).

Par ailleurs, on utilise de la voix numérisée pré-enregistrée comme voix pour le pseudo-pilote automatisé.

A l'instigation de la FAA, deux systèmes distincts de reconnaissance de la parole sont expérimentés pour le système ATCoach.

- Il y a d'une part un reconnaisseur de parole produit par Verbex qui opère en mode monolocuteur et qui utilise une syntaxe à états finis. Ce système accepte de la parole connectée. La taille du vocabulaire est d'environ 300 mots. La synthèse de parole est faite à partir de voix numérisée. Nous n'avons pas pu obtenir de résultats sur la performance de ce système pour le pseudo-pilote automatisé.

- Par ailleurs, il y a aussi un reconnaisseur de parole produit par Speech Systems Inc. (SSI) de Tarzana, Ca. Ce système utilise aussi une grammaire à états finis et est réputé être multilocuteur; il y a toutefois une possibilité d'adapter le modèle multilocuteur (on utilise le terme "profiler" le modèle) pour des usagers qui obtiendraient de mauvais résultats. Ce système est basé sur des réseaux neuroniques et fonctionne avec de la parole connectée. Le système utilisé jusqu'ici a été entraîné avec des mots tirés du vocabulaire courant : il n'y a eu aucun apprentissage du système avec le vocabulaire spécifique au contrôle aérien. On prévoit cependant faire cet apprentissage dans le cadre d'un projet de développement (en cours) réalisé pour le compte de la FAA. Le tableau suivant illustre une partie des résultats qui sont disponibles (13) :

Tableau I **Performance du reconnaisseur de parole SSI pour le système ATCoach**

Mode	% d'erreur mot	% d'erreur phrase
Multilocuteur	6,7 %	28,5 %
Multilocuteur avec profilage pour un locuteur	1,0 %	10 %

On constate que même si la performance de ce système multilocuteur peut être améliorée considérablement en utilisant le profilage pour un locuteur, il reste un taux d'erreur appréciable pour une phrase (10 %).

2.2.3 Système d'apprentissage de la phraséologie ATC

La société ITT a produit un système d'apprentissage de la phraséologie ATC qui comporte les caractéristiques suivantes (14) :

- Reconnaissance de parole monolocuteur en mots connectés, avec le reconnaisseur ITT VRS 1280. La plateforme est un PC 286 (AT), qui tourne sous DOS. Le langage de programmation est C.

- Utilisation d'une syntaxe à états finis "souple", qui permet de reconnaître des phrases comportant certaines constructions syntaxiques légales, ainsi que des constructions illégales. Les blocs sémantiques (indicatifs d'appel, autorisations, etc.) susceptibles d'être rencontrés sont inclus au même niveau dans la syntaxe. Une phrase peut donc être reconnue en dépit de l'inversion de ses constituants.

L'apprentissage consiste à comparer des phrases reconnues à des scénarios comportant les blocs sémantiques voulus. Un ensemble de règles, écrites en C, servent à comparer ce qui a été dit

avec ce qui aurait dû être dit. Si des erreurs sont détectées, l'élève en est informé.

2.3 Europe

Cette section a été élaborée à partir d'informations recueillies dans la littérature, de visites effectuées en France et en Angleterre et au cours d'une conférence organisée par l'Institut Eurocontrol à Luxembourg en octobre 1989. Au niveau européen, les deux intervenants principaux dans l'élaboration d'un pseudo-pilote automatique semblent être la France et l'Angleterre. On décrira donc l'état de la recherche appliquée dans ces deux pays.

2.3.1 France

Depuis plusieurs années, des chercheurs français étudient diverses tâches reliées au contrôle aérien. Entre autres, diverses études portant sur le langage utilisé par les contrôleurs aériens en situation réelles ont permis de dégager des caractéristiques importantes du langage utilisé par les contrôleurs (15). En s'appuyant sur certaines de ces études, le Laboratoire d'Informatique et de Mécanique pour les Sciences de l'Ingénieur (LIMSI), rattaché au Centre National de Recherche Scientifique, en collaboration avec le Centre d'Etudes de la Navigation Aérienne (CENA) et un partenaire industriel a élaboré un concept de pseudo-pilote automatique (16). Ce dernier est abordé comme un système complet, destiné à remplacer un humain. L'élève contrôleur est en interaction complète avec le système, de la même façon qu'il interagirait avec un pseudo-pilote humain. L'objectif du projet est donc de produire un système qui soit à la fois convivial et robuste. On reconnaît que le système doit composer avec les constatations suivantes :

- Il peut exister plusieurs formulations pour le même message.
- Il peut se produire des erreurs de la part du reconnaisseur de parole ou de la part de l'élève contrôleur.

L'approche utilisée au LIMSI intègre des connaissances

reliées à la tâche, au vocabulaire, à la phraséologie et à l'historique du dialogue sous forme d'une hiérarchie de schémas (17). La figure qui suit décrit sommairement le système (adapté de (18)) :

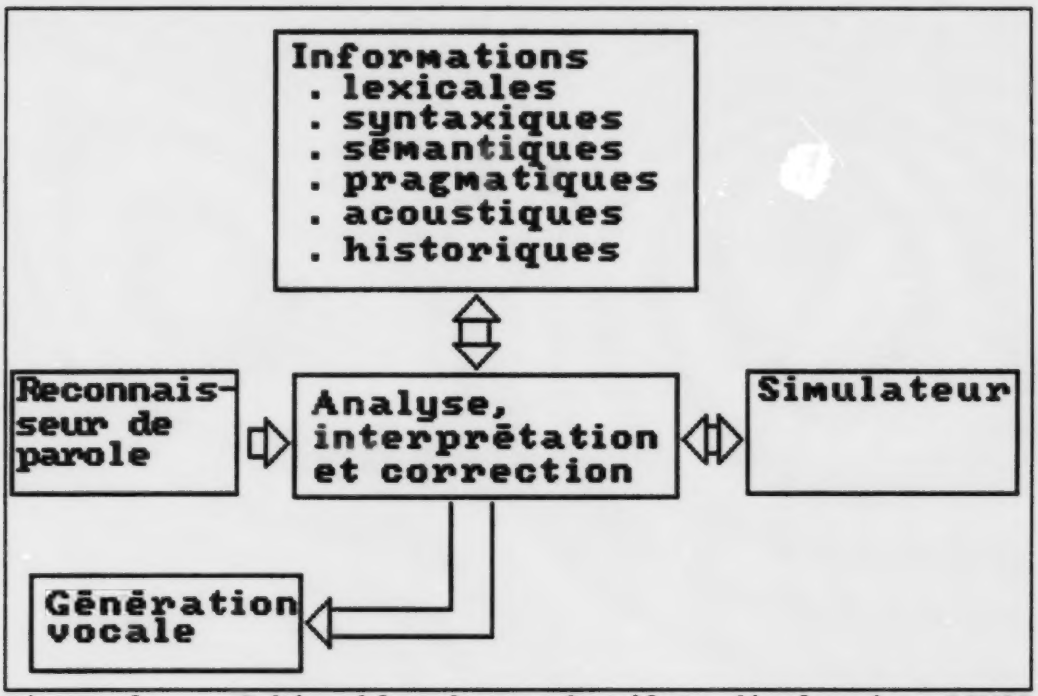


Figure 2 **Schéma-bloc du pseudo-pilote développé au LIMSI**

. L'élève contrôleur émet un message. Ce message est reconnu par un reconnaisseur de parole. Dans une version précédente du système, le reconnaisseur opérait en syntaxe libre, i.e. tous les mots du vocabulaire étaient susceptibles d'être reconnus. L'évolution du système vers des tâches plus complexes a nécessité l'utilisation d'une forme de syntaxe souple qui tout en améliorant les performances du système permet une grande liberté d'expression pour l'élève contrôleur. La syntaxe souple qui est utilisée évolue en fonction du dialogue et est appelée syntaxe évolutive binaire. Cette dernière consiste à définir pour chaque mot du vocabulaire une liste de mots pouvant le suivre. Une discussion plus complète sur l'évolution de la syntaxe se trouve dans (19).

- Ce message est associé à un schéma selon une stratégie de pairage entre les mots du message et les mots compris dans un schéma. On associe ensuite les éléments du message à des cases du schéma correspondant (instanciation).

- On procède alors à la détection et à la correction d'erreurs éventuelles. Un message doit être cohérent par rapport à l'historique du dialogue et par rapport à lui même (par exemple, un appareil ne peut monter à une altitude inférieure à son altitude actuelle). Diverses stratégies sont utilisées pour tenter de corriger une erreur sans avoir à faire appel à l'élève contrôleur. C'est seulement si la correction automatique a échoué que l'on demande à l'usager de répéter une partie ou la totalité du message.

- On met alors à jour l'historique du dialogue pour tenir compte du message et on prend l'action appropriée en réponse au message (par exemple, faire écho au message de l'élève contrôleur et changer de cap).

- Enfin, on tente de prédire de quel type sera le prochain message émis par l'élève contrôleur en utilisant un modèle de la tâche et un modèle de l'usager. Les prédictions sont utilisées pour modifier la syntaxe et ainsi faciliter la tâche du reconnaisseur de parole.

A son état d'avancement actuel, le système est matérialisé sous forme d'un démonstrateur qui possède les caractéristiques suivantes :

- Le reconnaisseur de parole est monolocuteur, et il reconnaît des mots connectés. On utilise actuellement le système Datavox développé par le LIMSI et commercialisé par Vecsys.

- Le modèle de langage utilisé est basé sur des schémas, auxquels sont associés les différents types de message susceptibles d'apparaître dans le dialogue.

- La voix du pseudo-pilote utilise la concaténation de mots. Le système incorpore un modèle de l'univers de

chaque avion et bâtit un historique du dialogue.

. La plateforme utilisée est un PC 386. La programmation est réalisée en Lisp et en C.

Selon les résultats disponibles (16), dans le cas d'un vocabulaire de 110 mots et en utilisant une syntaxe évolutive binaire, on obtient un taux de reconnaissance de 99.5%. Ce taux correspond à une compréhension sémantique correcte du message, en dépit des erreurs du reconnaisseur de parole.

2.3.2 Grande-Bretagne

L'intérêt pour l'application de la reconnaissance de la parole au domaine du contrôle aérien date de plusieurs années dans ce pays (20), (21). Même si on envisage pas d'applications à court terme des technologies vocales dans le milieu opérationnel, on veut évaluer et utiliser les technologies vocales dans des rôles pour lesquels on croit qu'elles conviennent actuellement. Ainsi, on a commandité une étude des aspects humains de la reconnaissance de la parole et de la phraséologie utilisée par des contrôleurs aériens dans des communications réelles (22).

Pour ce qui est de la réalisation d'un pseudo-pilote automatique, les Britanniques semblent avoir emprunté, du moins à court terme, une avenue différente des Français. Au lieu d'un système autonome destiné à la formation des élèves contrôleurs, les Britanniques ont développé un système hybride utilisant à la fois la voix et un écran tactile comme dispositif d'entrée. Ce système est utilisé pour des exercices auxquels participent des contrôleurs expérimentés. Ces derniers n'ont pas à faire l'apprentissage de la phraséologie ni des procédures de contrôle aérien : ils utilisent la voix comme canal d'entrée principal pour donner des instructions au pseudo-pilote automatique, quitte à utiliser un écran tactile si une commande présente des difficultés. Le système répond avec une voix synthétique (Dectalk).

Une version améliorée de ce pseudo-pilote est prévue pour 1990. On utilisera un reconnaisseur de parole (possiblement Macrospeak), mais on gardera quand même la possibilité d'utiliser une entrée manuelle (écran tactile). Une caractéristique intéressante de ce nouveau poste est que

l'acquisition des échantillons de voix nécessaires pour entraîner le reconnaisseur de parole se fera de façon transparente pour les utilisateurs. On utilisera des unités portatives servant à fournir un complément de formation aux contrôleurs actuellement en service pour recueillir les échantillons de voix. Les échantillons de voix d'un contrôleur donné seront chargés dans le reconnaisseur de parole du pseudo-pilote quand le contrôleur en question entreprendra une session de simulation.

2.4 Synthèse de la recherche appliquée actuelle

Cette revue de la recherche n'est pas exhaustive. En effet, une bonne partie de cette recherche est de nature très appliquée : elle est souvent effectuée par des sociétés commerciales qui ne publient pas leurs résultats. Il est donc possible que certains systèmes n'aient pas été mentionnés ici. Pour les systèmes qui ont été examinés jusqu'ici, on peut toutefois faire plusieurs constatations. D'abord, on note qu'il existe selon les diverses réalisations des différences importantes au niveau de :

- La situation de contrôle aérien pour laquelle on réalise un pseudo-pilote automatisé (contrôle de tour vs contrôle avec couverture radar).
- Le modèle de langage utilisé (syntaxe à états finis vs modèle basé sur des schémas).
- Le rôle du pseudo-pilote automatisé (simple simulateur vs conseiller dans l'exécution de la tâche).

On constate aussi qu'il n'y a pas eu d'évaluation complète de ces systèmes. On dispose en effet d'un certain nombre de résultats dans le cas de prototypes, mais comme aucun de ces systèmes n'est actuellement utilisé de façon régulière dans un centre de formation, il n'y a pas eu d'évaluation globale de leur performance et on ne peut juger si tous les objectifs visés ont été atteints. De plus, des discussions avec les concepteurs de certains des systèmes indiquent que plusieurs choix ont été faits de façon intuitive, sans examen des aspects ergonomiques. Par exemple, le choix de voix numérisée pré-enregistrée comme voix du pseudo-pilote automatisé est essentiellement basé sur une évaluation subjective de la qualité de la parole générée et

n'a pas fait l'objet d'une évaluation rigoureuse.

Selon ces constatations et suivant les objectifs qui ont été présentés précédemment, on est maintenant prêt à passer à la deuxième partie de ce projet : il s'agit d'élaborer un modèle sur mesure du langage utilisé par les élèves contrôleurs. Celà implique de franchir les étapes suivantes :

- Choisir pour quelle situation de contrôle aérien on veut modéliser le langage.
- Choisir quel formalisme on utilisera pour réaliser le modèle.
- Déterminer la démarche qu'on utilisera pour élaborer le modèle.
- Réaliser le modèle.

Le prochain chapitre amorce le processus en choisissant la situation de contrôle aérien pour laquelle on élaborera le modèle du langage.

3. CHOIX D'UNE SITUATION DE CONTRÔLE AÉRIEN POUR L'ÉLABORATION D'UN MODÈLE DE LANGAGE

Ce chapitre vise à choisir une situation de contrôle aérien qui soit favorable à l'élaboration d'un modèle sur mesure du langage utilisé par les élèves contrôleurs. On discute d'abord de façon générale de la phraséologie utilisée dans les différentes situations de contrôle aérien. On choisit ensuite une de ces situations comme objet de cette étude, et on la décrit de façon plus détaillée. Enfin, on donne un aperçu de l'environnement physique du poste qui a été étudié. Cette étape est préalable à l'analyse du langage des élèves contrôleurs. L'analyse du langage pour la situation de contrôle aérien choisie est présentée dans un chapitre ultérieur.

3.1 Discussion sur la phraséologie utilisée dans les différentes situations de contrôle aérien

De façon simpliste, on divise d'abord le contrôle aérien en contrôle de "vol à vue" et en contrôle de "vol aux instruments". On examine ensuite la phraséologie utilisée dans chacune de ces situations et on choisit la situation la plus favorable à l'élaboration d'un modèle de langage.

3.1.1 Contrôle de vol à vue

Le contrôle d'appareil en vol à vue se fait sans couverture radar : le contrôleur est en contact visuel direct avec l'appareil. Les directives transmises par le contrôleur au pilote font référence à des repères visuels facilement discernables par ce dernier. Il s'ensuit que la phraséologie est beaucoup moins standardisée en situation de vol à vue, comparativement à une situation de vol aux instruments. Des discussions avec des directeurs de cours à l'Institut de formation de Transports Canada confirment que dans certains cas, la phraséologie officielle ne suffit pas : il faut y suppléer avec du langage naturel. La modélisation du langage sous une forme utilisable par un reconnaisseur de parole est alors très

difficile. De plus, un pseudo-pilote automatisé pour le contrôle de vol à vue doit incorporer un environnement visuel représentant de façon réaliste ce que peut observer un contrôleur à partir d'une tour de contrôle. Cela augmente considérablement la complexité de réalisation du pseudo-pilote automatisé. Pour ces raisons, on décide de ne pas retenir cette situation pour le développement d'un modèle du langage.

3.1.2 Contrôle de vol aux instruments

D'après des discussions avec des directeurs de cours de l'Institut de formation de Transports Canada, la phraséologie utilisée dans le contrôle en vol aux instruments est très standardisée. De plus, on encourage à l'Institut un respect assez stricte de la phraséologie officielle. C'est donc ce genre de situation que l'on analysera pour voir comment on peut réaliser un pseudo-pilote automatisé. On doit aussi distinguer le contrôle en vol aux instruments selon qu'il s'effectue ou non avec couverture radar. L'examen de la documentation décrivant la phraséologie officielle (23) n'a pas fourni d'indications sur les avantages de choisir entre le contrôle aux instruments avec ou sans couverture radar. On choisit donc arbitrairement la situation de contrôle de vol aux instruments avec couverture radar.

3.2 Description de la situation considérée dans cette étude

Le type de contrôle effectué en vol aux instruments avec radar dépend de l'altitude et de la portion de l'espace considérées. La figure suivante montre une coupe d'un espace aérien fictif utilisé pour les exercices à l'Institut de formation de Transports Canada à Cornwall : il s'agit d'une partie de la zone de contrôle régionale de Midtown (24).

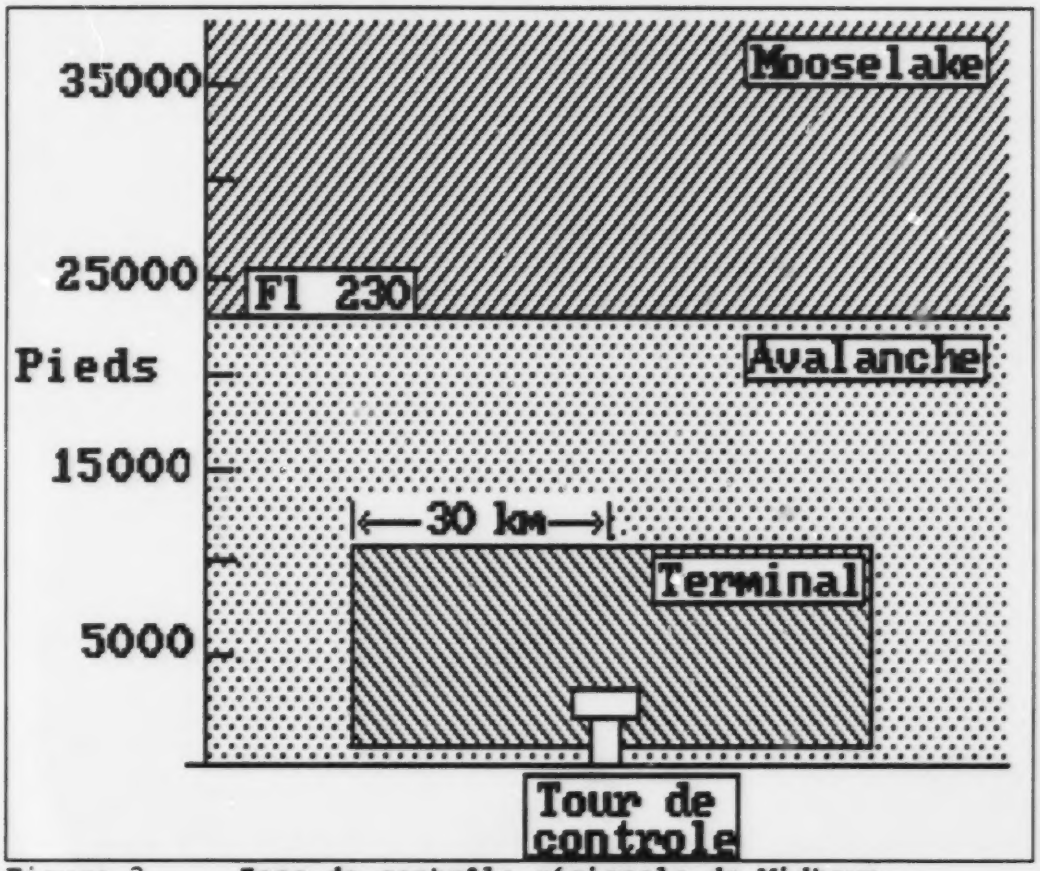


Figure 3 **Zone de contrôle régionale de Midtown**

L'échelle de gauche indique les altitudes. Par convention, les altitudes entre 0 et 18 000 pieds sont exprimées en milliers de pieds (par exemple, 11 000 pieds). Les altitudes égales ou supérieures sont indiquées en "niveaux de vol" : un niveau de vol correspond à un multiplicateur qui est appliqué à une hauteur de 100 pieds (par exemple, "niveau de vol 1 8 0" correspond à 180 * 100 pieds, soit 18 000 pieds).

On constate que la zone de contrôle régionale de Midtown est elle-même découpée en plusieurs zones. Dans le cas des exercices qui ont été enregistrés à l'Institut de formation de Transports Canada à Cornwall, on s'intéresse aux zones suivantes :

. Une zone de contrôle pour les altitudes élevées : la

zone Mooselake englobe toutes les altitudes à partir du niveau de vol 230.

- Une zone de contrôle pour les altitudes peu élevées : la zone Avalanche englobe toutes les altitudes à partir du sol jusqu'au niveau de vol 230.

- Une zone de contrôle de terminal : le terminal de Stilpond assure l'interface entre l'aéroport et les autres zones, pour les atterrissages et les décollages.

Les caractéristiques (fréquences radios à utiliser, routes, etc.) et les frontières géographiques de ces zones sont définies dans la documentation remise aux élèves contrôleurs. De plus, des procédures propres à chaque zone régionale prescrivent le protocole qui doit être suivi pour transférer un appareil à l'intérieur d'une zone régionale. Par exemple :

- Il y a un cap et une altitude préférentiels pour transférer un appareil entre la zone de haut niveau (Mooselake) et la zone de bas niveau (Avalanche).

- Le processus de transfert d'un appareil à une autre zone doit être enclenché à des moments précis lors du trajet d'un appareil vers une autre zone.

Il existe des mécanismes semblables entre les diverses zones régionales. Dans ce cas, les procédures sont énoncées dans des accords entre les zones régionales.

Pour choisir la portion de la tâche de contrôle en vol aux instruments qui sera retenue, on doit considérer plusieurs facteurs :

- La complexité de la tâche : on a avantage à tenter d'abord de modéliser une tâche simple.

- Le cheminement des élèves contrôleurs lors de leur apprentissage en simulation : les élèves contrôleurs ont une formation graduée qui les familiarise à tour de rôle avec toutes les procédures de contrôle utilisées dans les différentes zones.

Après discussion avec des directeurs de cours à l'Institut de formation, il semble que la tâche de contrôle en zone de

terminal soit la plus complète (i.e., elle fait intervenir le plus de procédures). Comme il s'agit aussi d'une tâche assez complexe, on adopte l'approche suivante :

- On examine une tâche simple, d'où on tire un modèle élémentaire mais complet. On effectue donc des enregistrements d'un exercice élémentaire qui se déroule dans une zone de contrôle en altitude peu élevée. Lors de cet exercice, les élèves contrôleurs en sont au début de leur formation en contrôle aux instruments.

- On tente ensuite d'adapter ce modèle à une tâche plus complexe, en l'occurrence le contrôle de terminal. Les élèves contrôleurs sont alors vers la fin de leur formation en contrôle aux instruments.

Cette approche progressive correspond ainsi au plan de cours utilisé à l'Institut de formation (voir la description du contenu du Segment Two, à l'annexe A).

3.3 Description du poste de l'élève contrôleur

3.3.1 Aspects physiques du poste

La figure suivante montre un croquis du poste de l'élève contrôleur. Le croquis n'est pas à l'échelle.

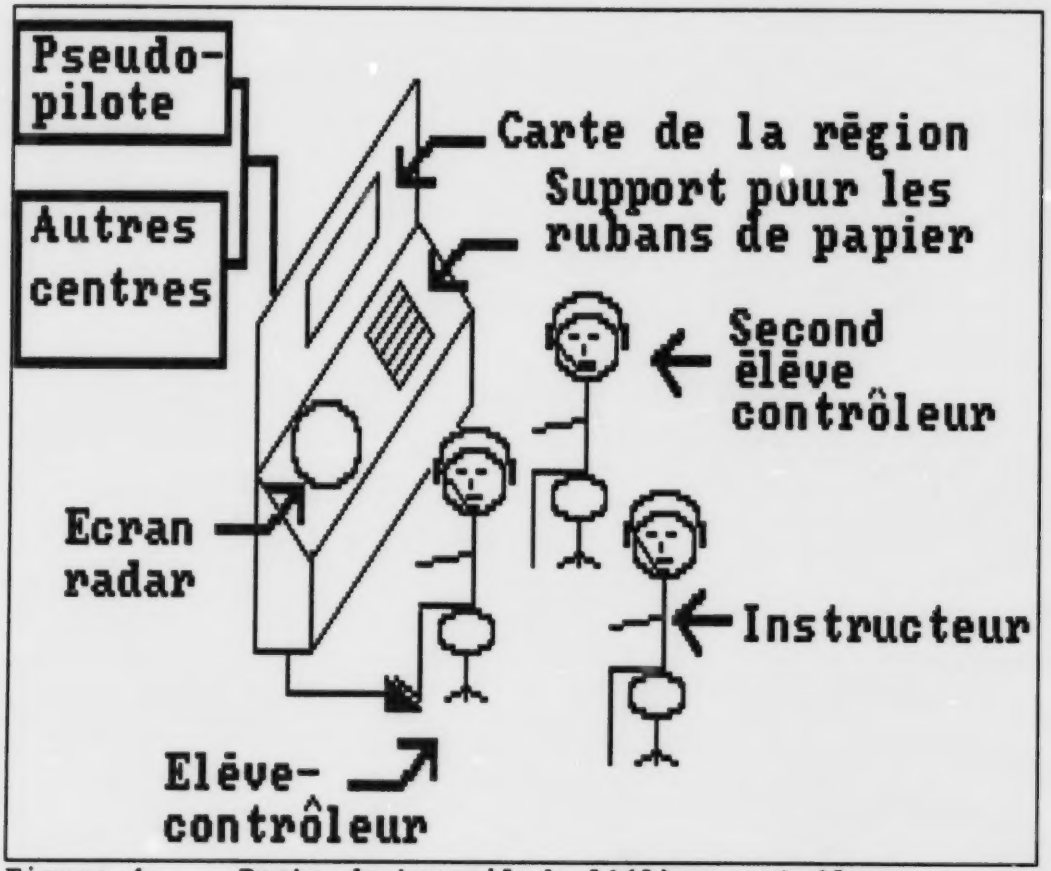


Figure 4 **Poste de travail de l'élève-contrôleur**

L'élève contrôleur et le pseudo-pilote sont situés physiquement dans des parties séparées de la salle de simulation.

3.3.2 Description générale des communications

Seules certaines des communications de l'élève contrôleur sont pertinentes pour cette étude. On peut classer ces communications comme suit :

- Communications avec l'instructeur qui est généralement situé directement derrière lui. L'élève contrôleur peut demander ou recevoir des conseils ou des

informations à l'instructeur. De plus, l'élève contrôleur peut aussi communiquer avec un second élève contrôleur situé au même poste. Ce second élève contrôleur assure, entre autres, la coordination avec d'autres centres et n'intervient pas dans les communications avec le pseudo-pilote. Les communications entre les deux élèves contrôleurs et l'instructeur se font sans l'intermédiaire d'aucun appareil.

. Communications avec d'autres centres. Ces communications se font par l'intermédiaire du même casque que celui utilisé pour communiquer avec les pilotes. Le même canal est aussi utilisé, à une variante près : un interrupteur manuel sert à aiguiller l'appel vers les autres centres. Dans une situation de contrôle aérien réel, le pilote n'entend pas ces communications.

. Communications avec le pseudo-pilote. Ces communications se font par l'intermédiaire d'un casque. Un interrupteur dédié, commandé au pied, permet à l'élève contrôleur de parler au pseudo-pilote.

3.3.3 Observations sur le déroulement de la tâche

Un exercice dure de 35 à 40 minutes. Avant le début de l'exercice vocal proprement dit, un instructeur explique au moyen d'acétates les différentes situations qui pourront se présenter. Les élèves contrôleurs prennent alors position et l'exercice vocal commence.

On note un certain nombre de phénomènes susceptibles d'avoir une influence sur un système de reconnaissance de la parole :

. Lors de l'exercice, l'élève contrôleur chargé du poste répond aux requêtes des pilotes ou initie les échanges. On observe que le port du casque d'écoute varie quelque peu selon les élèves contrôleurs : certains portent leur casque d'écoute de façon "correcte" et ne semblent pas avoir tendance à en varier la position alors que d'autres vont changer sa

position lors de l'exercice.

- Il y a quelques fois des conversations près du contrôleur lorsque ce dernier est en interaction avec le pilote.

- Les contrôleurs sont contraints à utiliser des rubans de papier pour la gestion de l'information (les "strips"). Ces rubans sont insérés dans des supports, et les supports sont déplacés, manipulés, etc. La manipulation de ces supports est assez bruyante : les bruits sont courts et prononcés. L'intégration de technologies vocales devrait tenir compte de ceci.

- Lors des observations, il est apparu qu'un contrôleur parlait à un seul pseudo-pilote qui jouait le rôle de plusieurs pilotes. Le contrôleur entend donc toujours la même voix. De plus, le pseudo-pilote peut aussi jouer le rôle du contrôleur d'un autre centre, lequel est en communication avec l'élève contrôleur.

Il reste une étape à franchir avant d'élaborer la démarche de modélisation du langage. En effet, on doit choisir quel formalisme on utilisera pour représenter (ou modéliser) le langage. C'est ce qu'on fait au chapitre suivant.

4. CHOIX DU MODÈLE DE REPRÉSENTATION DU LANGAGE CONTRÔLEUR-PILOTE

Ce chapitre vise à choisir un modèle de représentation du langage utilisé lors des échanges contrôleur-pilote. On effectue le choix en analysant les modèles qui ont été utilisés jusqu'ici dans la conception de pseudo-pilotes automatisés. On entend par modèle de représentation du langage un formalisme décrivant :

- Le lexique effectif du langage opératif.
- La structure du langage opératif.

Ce modèle est nécessaire pour permettre la reconnaissance de phrases par un système de reconnaissance de la parole, en fournissant le lexique et éventuellement la syntaxe du langage. De plus, le développement du modèle permettra d'améliorer la compréhension de la tâche. Enfin, le modèle sera aussi utile pour élaborer les exigences requises à l'égard du système de reconnaissance et de génération de parole.

4.1 Analyse de modèles existants

Les travaux effectués pour l'application de la reconnaissance de la parole dans le contrôle aérien ont utilisé jusqu'ici deux approches pour modéliser le langage opératif utilisé en contrôle aérien (8) (13 ,14) (15) (18) (25). Chacune de ces approches est adaptée à une école de pensée différente. En effet, on peut avoir :

- Une philosophie "stricte" : on est alors partisan d'une utilisation rigide de la syntaxe. On restreint alors le nombre de formulations permises pour un même signifié.
- Une philosophie "souple" : on permet diverses formulations pour un même signifié et on tolère un certain nombre de déviations autour d'une syntaxe donnée, par exemple des abréviations, des inversions, etc.

On examine ici les modèles correspondants à chacune de ces philosophies.

4.1.1 Représentation sous forme d'un arbre syntaxique.

Cette représentation englobe le lexique et la syntaxe du langage sous la forme d'une structure arborescente. Une phrase composée de mots appartenant au lexique et syntaxiquement valide se retrouve comme une succession d'embranchements dans l'arbre. La figure suivante illustre l'utilisation d'un modèle (hypothétique) de ce type pour reconnaître la phrase "turn left heading 355".

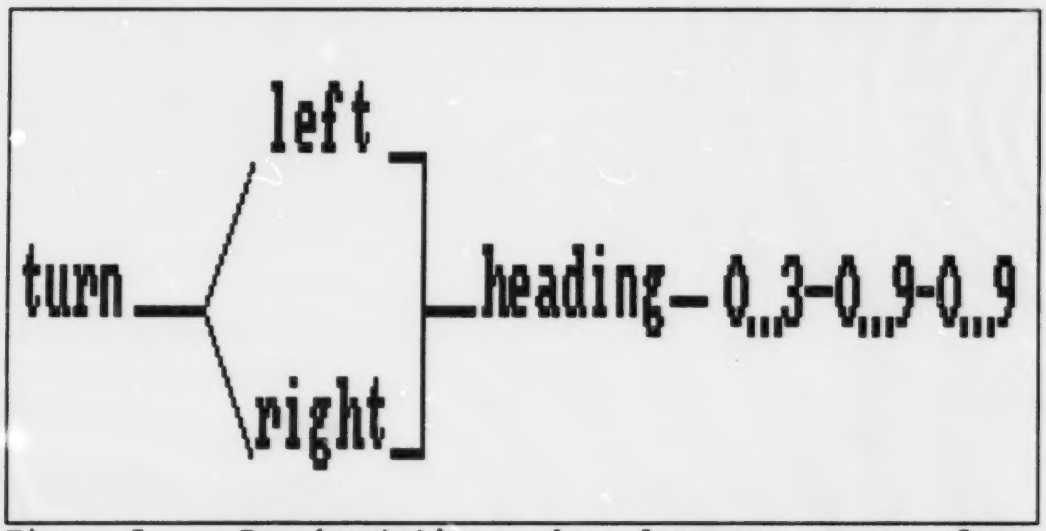


Figure 5 **Représentation du langage sous forme d'arborescence**

Cette approche permet de réduire, parmi l'univers des mots du vocabulaire, le nombre de mots susceptible de suivre le dernier mot reconnu; on peut ainsi sensiblement améliorer la performance des reconnaisseurs de parole. Cette représentation est celle qui a dominée depuis l'introduction des systèmes de reconnaissance de la parole. En première approximation, comme ce modèle recouvre l'ensemble du lexique et de la syntaxe permise, on devrait pouvoir décrire tous les échanges vocaux légaux susceptibles de se produire. Ce pourrait être le cas si la production des phrases était toujours parfaitement contrôlée dans une tâche de contrôle aérien. L'analyse de plusieurs

corpus tirés de conversations entre pilotes et contrôleurs dans des situations réelles montre que (15) (22) :

- On retrouve diverses formulations (soit officielles, soit consacrées par l'usage) pour le même signifié.

- On retrouve aussi un certain nombre "d'erreurs" au sens syntaxique, soient des pauses, des hésitations ("euh", "err", etc.), des additions ("bonjour"), des reprises, etc.

Si on veut utiliser une structure arborescente comme modèle du langage tout en conservant souplesse et transparence au dialogue, il faut d'y inclure plusieurs formulations différentes pour le même signifié (donc ajouter des branches à l'arbre ou enrichir des embranchements existants). Cela a toutefois deux conséquences indésirables :

- Cette approche augmente le facteur de branchement de la structure : le facteur de branchement correspond au nombre de branches possibles suivant un mot donné. Si le facteur de branchement est élevé, il est plus difficile pour le reconnaisseur de parole de choisir le bon branchement.

- Cette approche peut devenir rapidement lourde en termes d'apprentissage pour le reconnaisseur de parole.

Même si ces deux difficultés étaient surmontées, d'autres problèmes sont susceptibles de se poser. Par exemple, l'emploi d'un mot inconnu peut conduire à ne pas reconnaître une phrase parfaitement légale par ailleurs.

Pour ce qui est des erreurs (pauses, hésitations, reprises, etc.), on peut tenter de contourner le problème en tentant de prédire où certaines erreurs sont susceptibles de se produire et d'ajouter des éléments correspondants (i.e. des pauses, des "euh", etc.) dans le modèle pour y pallier. On rencontre alors la difficulté suivante :

- Il est difficile de prévoir exactement où se produiront les erreurs de dialogue. A cause de la rigidité de la représentation arborescente, il est nécessaire de pouvoir modéliser exactement où se produiront chacune des erreurs.

L'ajout de mots dans la syntaxe (par exemple, "correction") pour tenir compte de la présence d'erreurs conduit aussi à l'alourdissement de l'arbre syntaxique. De plus, il n'est pas simple d'implanter un mécanisme pouvant tenir compte de toutes les possibilités de reprises et de corrections de la part du locuteur, avec ce genre de modèle.

4.1.2 Représentation sous forme de schémas

Cette approche exploite le fait que les échanges verbaux, dans le contexte du contrôle aérien, peuvent être associés à un langage opératif et classés dans un nombre restreint de catégories. Ces catégories sont alors associées à des schémas. Un schéma est une structure de données qui décrit une situation stéréotypée (26) (ici des types d'échanges verbaux) et consiste en un ensemble de cases. Ces cases peuvent contenir des références à d'autres schémas, des conditions, des valeurs par défaut ou des procédures. Un tel modèle de représentation de connaissance compte généralement plusieurs schémas (en fait au moins un schéma par catégorie de message dans le langage). Dans le cadre du contrôle aérien, cette approche a été utilisée par (15). La figure suivante illustre l'utilisation d'un modèle de ce type pour reconnaître la phrase "turn left heading 355".

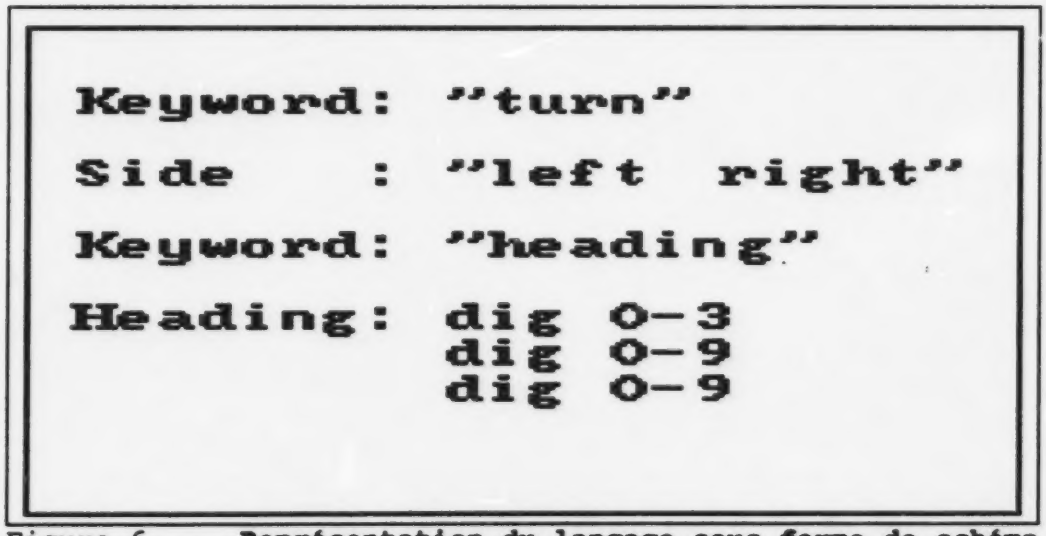


Figure 6 **Représentation du langage sous forme de schéma**

On constate que ce mode de représentation est fondamentalement différent du modèle précédent (arbre syntaxique). Alors que ce dernier analyse une phrase en utilisant des contraintes syntaxiques, le modèle basé sur des schémas tente plutôt d'abord d'associer une communication à un contexte donné (dans la figure, on évoque un schéma relatif à une directive de changement de cap). Une fois le contexte identifié, on procède au pairage entre les éléments du contexte et le contenu du message (on tente donc d'identifier le coté, "left" ou "right", ainsi que le cap, "heading 355"). Cette représentation est donc orientée sur l'aspect sémantique du message; on ne fixe pas à priori dans quel ordre on doit retrouver les différents éléments dans la phrase. Par ailleurs, elle permet aussi d'utiliser une forme de syntaxe "locale" (i.e. on reconnaît la présence de noyaux syntaxiques à l'intérieur d'une phrase, sans que la localisation de ces noyaux soit fixée de façon rigide). L'exploitation du modèle pour la compréhension du message compte donc essentiellement deux étapes :

- D'abord, reconnaître à quelle catégorie appartient le message.
- Ensuite, instancier le contenu du message par rapport à ce qu'on s'attend à y retrouver.

L'utilisation de la représentation par schéma permet de résoudre de façon élégante les difficultés posées par la présence de plusieurs formulations pour un même signifié. En effet, l'actualisation du schéma permet un ordre différent des éléments clés du message. De plus, les "erreurs" sont simplement acceptées comme ce qu'elles sont : des événements susceptibles de se produire dans certaines situations mais difficiles à prévoir avec précision. On tire alors avantage du fait que certaines des cases du schéma sont facultatives.

Par ailleurs, une modélisation du langage utilisant des schémas s'intègre naturellement dans un système de pseudo-pilote automatisé utilisant une approche déclarative. On peut utiliser un formalisme uniforme orienté-objet pour représenter des connaissances sur l'univers de la tâche (zone de contrôle, appareils, et.) et rendre le pseudo-pilote automatisé plus intelligent.

Cette représentation présente aussi certains inconvénients. Entre autres :

- Pour pouvoir utiliser ce formalisme, il faut que le langage à modéliser s'y prête. Il faut que les catégories de messages ayant une signification différente soient suffisamment distincts l'un de l'autre pour qu'un message puisse être classifié sans ambiguité dans la bonne catégorie.

- L'apprentissage de certains reconnaisseurs de parole tente de prendre en compte la coarticulation entre différents éléments du lexique. Si on permet à tous les mots de suivre n'importe quel autre mot (i.e. si on opère sans syntaxe), on peut rendre l'apprentissage du reconnaisseur de parole problématique. Il peut être bénéfique d'utiliser une forme de syntaxe "souple" qui permette d'améliorer la reconnaissance de certains éléments tout en facilitant l'apprentissage du reconnaisseur de parole.

On peut aussi tenter de générer une syntaxe évolutive, dépendant du contexte et de l'état du dialogue à un moment donné. Cette méthode est d'ailleurs utilisée dans un prototype actuel (19).

A un autre niveau, on peut restreindre le choix du schéma parmi un nombre restreint de schémas, plausibles dans la situation actuelle. Ceci revient à créer une "syntaxe de schémas" plus ou moins rigide, ou encore un script. Un script est une structure qui décrit une séquence appropriée d'événements dans un contexte particulier (27).

4.2 Choix du modèle de représentation du langage

Le choix du modèle de représentation dépend de plusieurs facteurs, et en particulier de ce que l'on souscrive à la philosophie "stricte" ou à la philosophie "souple". Le choix dépend aussi du niveau de la formation qu'on prévoit fournir au moyen du pseudo-pilote automatisé. La formation d'élèves contrôleurs tout à fait novices peut probablement tirer avantage d'une phraséologie très stricte, mais on peut croire que pour des élèves plus avancés, on ait intérêt à accepter une forme d'expression plus riche (comme c'est d'ailleurs actuellement le cas avec un pseudo-pilote humain).

Dans le cadre de cette étude, on adopte le modèle de représentation basé sur des schémas. On croit qu'on pourra obtenir un modèle plus souple, quitte à imposer des contraintes syntaxiques additionnelles plus tard si l'on désire une phraséologie très stricte. De plus, il nous semble que l'utilisation de techniques tirées de l'Intelligence Artificielle s'imposera tôt ou tard dans une application de pseudo-pilote automatisé. L'utilisation de schémas pour modéliser le langage permettra d'utiliser plus facilement ces techniques.

On est maintenant en mesure d'élaborer la démarche qui sera utilisée pour modéliser le langage utilisé par les élèves contrôleurs, pour la situation de contrôle aérien qui a été retenue.

5. DÉMARCHE DE LA MODÉLISATION DU LANGAGE OPÉRATIF

Cette section explique la démarche utilisée pour créer un modèle sur mesure du dialogue opérationnel utilisé par les élèves contrôleurs. On a convenu de mener l'étude pour des dialogues en anglais. Le modèle doit permettre de décrire les échanges entre les élèves contrôleurs et les pseudo-pilotes d'une façon qui soit utilisable par un pseudo-pilote automatisé. On a le choix entre modéliser soit la phraséologie officielle préconisée dans (23), soit la phraséologie réellement utilisée par les élèves contrôleurs. On choisit de modéliser le langage réellement utilisé par les élèves contrôleurs pour les raisons suivantes :

- On a mentionné au chapitre 4 qu'il y a dans les faits des différences entre la phraséologie officielle et celle qui est réellement utilisée par les contrôleurs. Un pseudo-pilote automatisé utilisant un modèle du langage réel sera plus "transparent" pour l'élève contrôleur; l'exercice sera vraisemblablement plus proche de la réalité.

- Un modèle est une simplification de la réalité. Même si on tente de modéliser la phraséologie réelle, il est douteux qu'on parvienne à en capturer toute la richesse d'expression. En ce sens, même un modèle de la phraséologie réellement utilisée est plus restrictif que la phraséologie réelle. Il s'agit donc d'un moyen terme entre la phraséologie officielle et la phraséologie réellement utilisée.

Comme le modèle de langage doit être utilisé pour comprendre le sens des directives des élèves contrôleurs, l'analyse du langage portera en priorité sur les messages provenant de ces derniers. On considère que les réponses du pseudo-pilote seront, dans la plupart des cas, des conséquences directes des directives reçues; l'analyse de ces réponses est donc secondaire ici.

5.1 Dialogue cible

Comme on l'a mentionné au chapitre 3, la complexité appréhendée du développement du modèle suggère de tenter d'abord de modéliser une situation simple. On modélisera d'abord des exercices de contrôle en altitude peu élevée. Par la suite, on verra comment on peut modifier le modèle obtenu pour l'adapter à une situation plus complexe, i.e. le contrôle de terminal.

5.2 Description de la démarche de développement du modèle

Au chapitre précédent, on a choisi un modèle de représentation du langage basé sur des schémas. Pour élaborer ce genre de modèle, on aura besoin des éléments suivants :

- Le vocabulaire des messages : le vocabulaire (ou lexique) que doit reconnaître le pseudo-pilote automatisé pour pouvoir reconnaître et interpréter les messages.

- Les catégories de messages : chacune des catégories de messages correspondra à un schéma.

- La structure des messages : chaque schéma compte un certain nombre de cases qui correspondent à des noyaux sémantiques (entités porteuses de sens). Ces cases (ou noyaux sémantiques) peuvent aussi correspondre à des noyaux syntaxiques (par exemple la formulation d'une altitude). Dans certains cas, des limites (minimum ou maximum), des méthodes de calcul, etc., peuvent être associées à ces cases.

On choisit de procéder en deux étapes pour obtenir ces éléments. On fait d'abord une analyse sommaire du dialogue d'où on tire une partie du modèle du langage. On complète alors ce modèle préliminaire lors d'une seconde étape. Les sections qui suivent décrivent ces deux étapes.

5.2.1 Description de la première étape du développement du modèle de langage

Cette première étape vise à :

- Obtenir le vocabulaire utilisé.
- Obtenir l'ensemble des schémas préliminaires qui représentent les catégories de messages qui sont échangés entre l'élève contrôleur et le pseudo-pilote.
- Définir le contenu des schémas préliminaires, i.e. déterminer les cases appartenant à chaque schéma. Ces cases reflètent les éléments sémantiques de chaque message (i.e. les entités porteuses de sens). De plus, lors d'une implantation éventuelle du modèle, le contenu des cases peut fournir des îlots de syntaxe locale.

La figure suivante décrit de façon schématique la première étape du développement du modèle de langage.

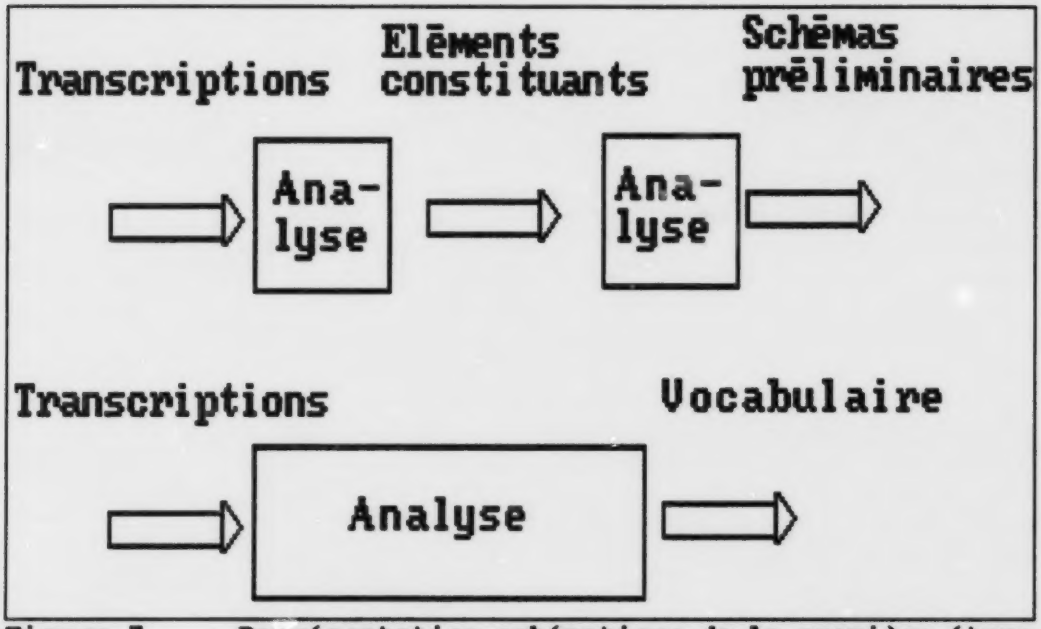


Figure 7 **Représentation schématique de la première étape du développement du modèle de langage**

D'une part, on examine chacun des messages provenant de l'élève contrôleur et on en identifie les éléments constituants (indicatifs d'appel, altitudes, etc.). On classe alors ensemble ceux qui présentent les mêmes traits (mêmes mots-clés, mêmes éléments constituants, longueur sensiblement équivalente). On procède ainsi pour chacune des transcriptions. On laisse toutefois de côté les situations comportant des erreurs de dialogue : ces cas seront traitées ultérieurement. On consolide les résultats de chacune des transcriptions, en procédant au besoin aux ajustements requis.

Par ailleurs, on établit aussi le vocabulaire des messages provenant de l'élève contrôleur pour chacune des transcriptions. On consolide ensuite les résultats.

L'élaboration des schémas préliminaires et l'établissement du vocabulaire peuvent être menés indépendamment l'un de l'autre.

5.2.2 Description de la deuxième étape du développement du modèle de langage

Le modèle préliminaire obtenu à l'étape précédente est incomplet : il manque d'une part les mécanismes d'instanciation de certaines cases des schémas, et d'autre part les mécanismes de gestion des erreurs. De plus, on veut déterminer s'il est possible de prédire quels schémas seront évoqués lors d'un exercice. Cette dernière étape du processus complète donc le modèle préliminaire. On doit atteindre les objectifs suivants :

- Dégager des mécanismes d'instanciation pour divers champs des schémas.
- Examiner sommairement les mécanismes de gestion des erreurs de dialogue.
- Dégager les règles d'évocation des schémas successifs.

Pour atteindre ces objectifs, on élabore d'abord des règles d'instanciation des schémas qui prennent avantage de toutes les ressources disponibles. En effet, comme le dialogue a lieu dans le cadre d'un exercice bien défini, la connaissance du "micromonde" du contrôleur et de celui du pilote peut être mise à contribution pour pallier les limitations de la reconnaissance de la parole. On peut donc enrichir les schémas de sources d'informations secondaires ou de valeurs par défaut qui sont utilisées quand le dialogue lui-même ne suffit pas.

Conjointement avec ce processus, on détermine comment les erreurs sont gérées lors du dialogue. On dégage alors des règles qui servent à modéliser cette portion du dialogue.

Enfin, on essaie de déterminer quelle est la succession des schémas évoqués lors de la réalisation d'un scénario donné. On dégage alors des règles pour aider à restreindre le nombre de schémas qui doivent être évoqués lors de l'analyse d'un message. On obtient alors un modèle fidèle du dialogue opératif. La figure suivante décrit schématiquement cette étape.

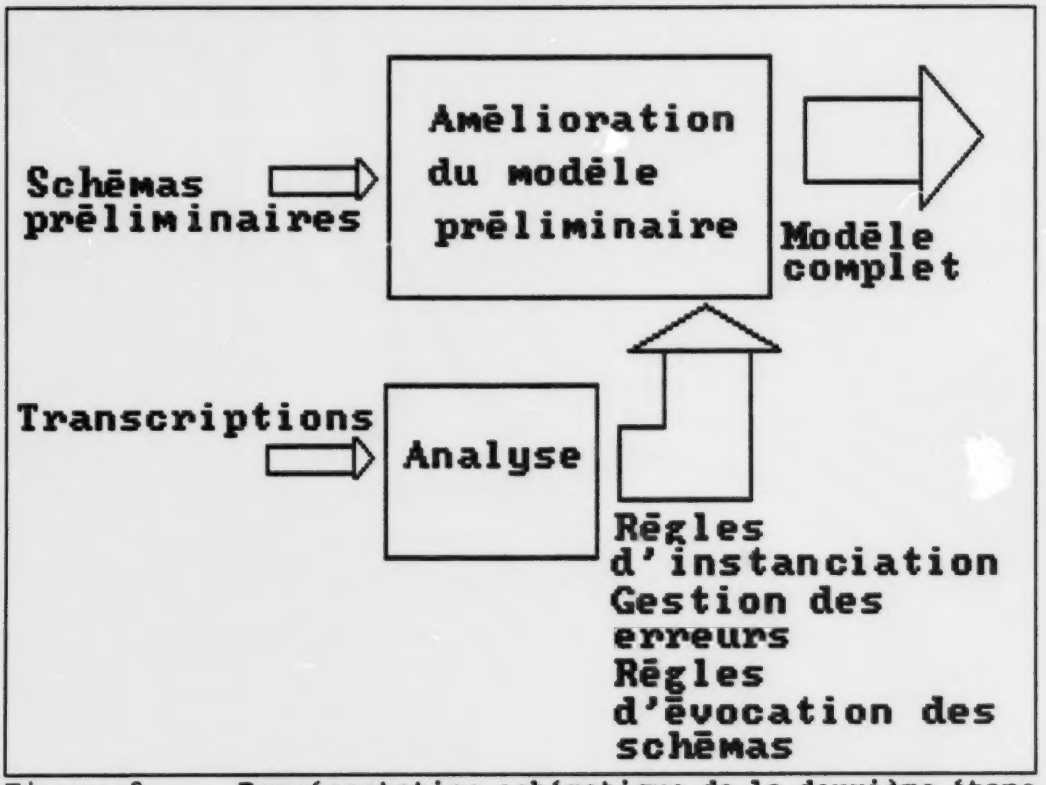


Figure 8 **Représentation schématique de la deuxième étape du développement du modèle de langage**

On peut maintenant élaborer le modèle proprement dit. Les deux chapitres qui suivent décrivent donc l'application de cette méthode à la modélisation du langage utilisé dans des exercices simples en altitude peu élevée.

6. PREMIERE ÉTAPE DU DÉVELOPPEMENT DU MODÈLE POUR DES EXERCICES SIMPLES

Cette section utilise la méthode présentée à la section précédente pour analyser des transcriptions de dialogues opératifs réels. Comme on l'a mentionné au chapitre 3, on a choisi d'analyser d'abord un exercice simple de contrôle radar en vol aux instruments en altitude peu élevée, pour lequel on a réalisé deux enregistrements à l'Institut de formation de Transports Canada, à Cornwall, Ontario. Chaque enregistrement dure environ 35 minutes et correspond à l'accomplissement d'un exercice par un élève contrôleur différent. On a transcrit chacun de ces enregistrements sous une forme orthographique la plus fidèle possible. Les transcriptions sont appelées 1a et 2a et se trouvent à l'annexe B. On avait prévu utiliser quatre enregistrements (donc quatre transcriptions), mais un problème technique a entraîné l'élimination de deux des enregistrements. Malgré la taille réduite de cet échantillon, on compte pouvoir obtenir un modèle préliminaire du langage opératif; on évaluera dans une phase ultérieure si ce modèle est valide, compte tenu de la taille réduite du corpus. Pour faciliter l'élaboration du modèle, on pose les contraintes suivantes :

. Les messages (ensembles cohérents d'information appartenant à une catégorie donnée) ne sont pas concaténés lors d'une transmission (communication radio).

.. Une transmission s'adresse à un seul destinataire.

On lèvera ces contraintes ultérieurement.

6.1 Examen de la longueur des transmissions

On examine ici la distribution de la longueur des transmissions émises par les élèves contrôleurs. Cette information sera utile pour donner une idée de la complexité du traitement qu'aura à accomplir le reconnaisseur de parole. La figure qui suit montre cette distribution, pour les exercices simples :

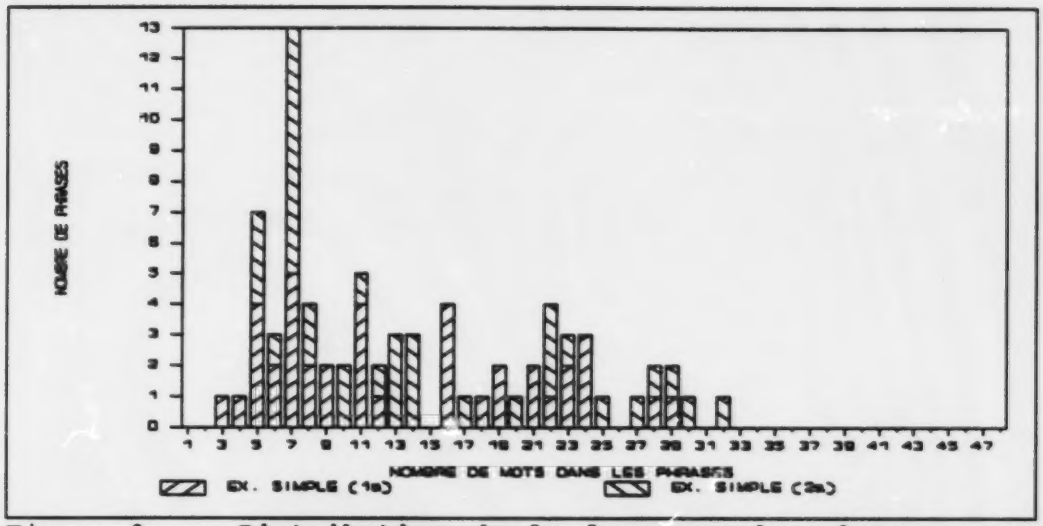


Figure 9 **Distribution de la longueur des phrases pour les exercices simples**

On constate que même dans le cas des exercices simples, on peut avoir des phrases pouvant compter jusqu'à 32 mots. Etant donnée la taille de l'échantillon, on peut considérer qu'il s'agit d'une borne inférieure, i.e. que la longueur maximale d'une phrase pour ce genre d'exercice est d'au moins 32 mots.

6.2 Obtention du vocabulaire

Dans cette section, on établit d'abord un ordre de grandeur de la taille du vocabulaire requis pour un exercice de ce genre. On détermine aussi, de façon préliminaire, quelle est la composition de ce vocabulaire. On sait que certaines classes d'éléments du vocabulaire vont dépendre du scénario de l'exercice : par exemple, les noms des autres centres de contrôle avec lesquels l'élève contrôleur doit faire affaire ou les noms de différentes lignes aériennes utilisées dans les exercices. De plus, le code d'épellation en radio-téléphonie de l'OACI (28) (alpha..zulu) et les nombres (avec diverses formulations) sont susceptibles d'être utilisés, mais leur utilisation peut varier selon le contrôleur (par exemple, le contrôleur peut dire "one hundred" au lieu de "one zero zero"). Comme on désire trouver un vocabulaire qui soit robuste, i.e. qui soit assez stable d'un exercice à l'autre, on désire

utiliser une liste stable de noms de centres de contrôle, de lignes aériennes, d'éléments du code d'épellation en radiotéléphonie et de nombres. On en tiendra compte dans l'obtention du vocabulaire requis.

Le tableau qui suit montre la taille du vocabulaire pour les transcriptions. La liste complète des mots de vocabulaire de chaque transcription (exercices simples) se trouve à l'annexe C.

Tableau II **Résultats de l'évaluation de la taille du vocabulaire**

	Transcription 1a	Transcription 2a
Nombre total de mots utilisés	74	67
Références à des centres de contrôle	5	4
Références à des lignes aériennes	7	7
Références au code d'épellation de l'OACI	4	4
Références à des chiffres	20	14
Vocabulaire opératif net (incomplet)	37	37

Chaque colonne du tableau correspond à une transcription. Pour

chaque transcription, on retrouve la taille du vocabulaire total. On soustrait les références à des centres de contrôle, des lignes aériennes, des éléments du code d'épellation en radio-téléphonie et des chiffres. On obtient alors le vocabulaire opératif net auquel il manque maintenant certains éléments.

Pour que l'on puisse établir un vocabulaire opératif, il faut que les différents usagers emploient non seulement un nombre restreint de mots, mais ils doivent aussi utiliser à peu de choses près les mêmes mots. Pour voir si c'est le cas, on combine les vocabulaires d'usage général nets de chacune des transcriptions (i.e. 1a et 2a). On obtient alors un vocabulaire opératif net combiné de 45 mots. Cela montre bien que la majorité des mots utilisés sont communs aux deux élèves contrôleurs. On peut alors supposer que pour le même genre de tâche, la taille du vocabulaire d'usage général utilisé par plusieurs élèves contrôleurs serait du même ordre.

Comme on l'a mentionné, pour obtenir un vocabulaire opératif robuste, on doit tenir compte d'un nombre défini de centres de contrôle, de lignes aériennes, d'éléments du code d'épellation en radio-téléphonie et de nombres. En utilisant un nombre standard de ces éléments, le vocabulaire devient relativement indépendant du détail des exercices (par exemple, les indicatifs d'appel peuvent différer largement d'un exercice à un autre). De plus, on peut ainsi contourner certaines différences individuelles entre les contrôleurs (par exemple, on peut utiliser "one zero zero" au même titre que "one hundred"). Le tableau qui suit montre comment on évalue le nombre requis de ces divers éléments.

Tableau III Evaluation du nombre d'éléments extrinsèques

Elément standard	Mode de calcul et source des données	Résultat
Noms de centres de contrôles	Il y a 8 zones pouvant être mentionnées lors des communications ("Study Notes")	8
Noms de lignes aériennes	Après consultation des transcriptions, on constate que 10 noms semblent suffire	10
Eléments de l'alphabet phonétique	(alpha bravo ... zulu)	26
Nombres et variantes	(zero one ... niner nine) (ten eleven..... nineteen) (twenty thirty ... ninety) (hundred thousand)	11 10 8 2
Directions	(north south east west)	4
Total		79

La dernière étape pour obtenir la taille du vocabulaire opératif requis est de faire la somme du vocabulaire combiné pour chaque type de transcriptions avec le nombre d'éléments standards. C'est ce que montre le tableau suivant.

Tableau IV Taille du vocabulaire opératif

Taille du vocabulaire d'usage général	Nombre d'éléments standards	Taille totale du vocabulaire opératif
45 mots	79 mots	124 mots

En gardant une marge de sécurité qu'on fixe arbitrairement à 25%, la taille totale du vocabulaire opératif est de 155 mots (124 mots * 1.25) pour ce type d'exercice. Cette marge servira à absorber les différences d'utilisation de vocabulaire entre les élèves contrôleurs et permettra aussi de tenir compte d'autres éléments (par exemple les noms de balise). Comme on l'a mentionné, il s'agit d'un ordre de grandeur et non d'une valeur exacte.

6.3 Élaboration des schémas préliminaires

Cette section montre comment sont élaborés les schémas préliminaires permettant de modéliser le dialogue. On parcourt d'abord la transcription de chacun des exercices en créant pour chaque message que l'on rencontre un schéma temporaire identifié par un nom. Lorsque l'on rencontre un message qui ressemble (mêmes mots clés, mêmes éléments constituants, longueur sensiblement équivalente) à un autre message pour lequel un schéma temporaire a déjà été généré, on considère que ce message appartient au même schéma. Le tableau qui suit montre un exemple de ce processus dans le cas du schéma le plus simple parmi ceux qui ont été élaborés. On peut trouver une liste complète des schémas individuels et des messages ayant servi à les élaborer à l'annexe D.

Tableau V Exemple d'élaboration d'un schéma

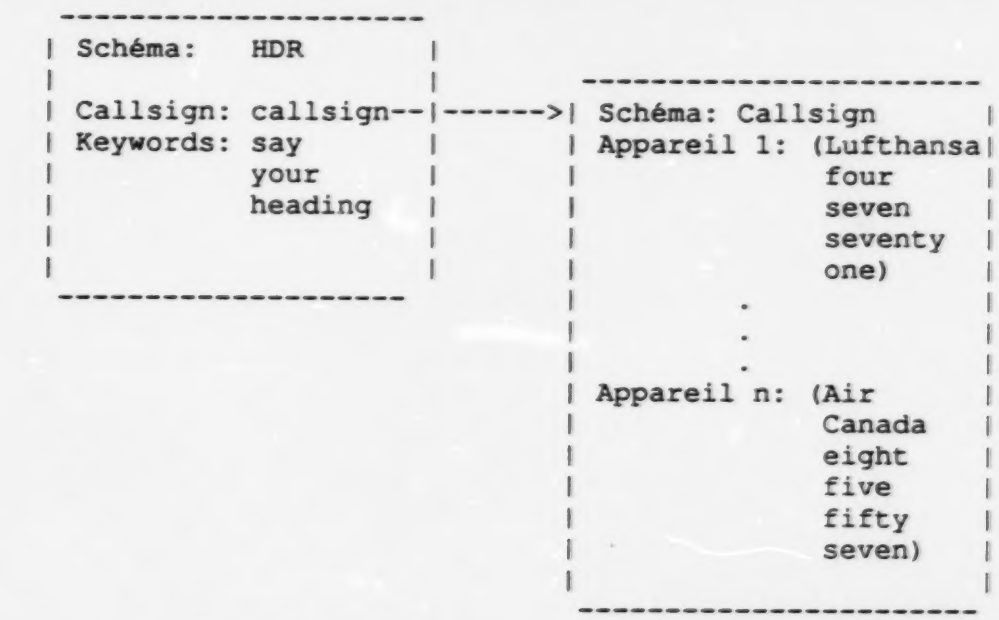
SCHEMA: HDR

Trans.no.	de msg	Contenu du message						
1a	1	november	nine hundred	papa	bravo	say	your	heading
	31	lufthansa	four seven	one		say	your	heading
	67	alitalia	six sixty	five		say	your	heading
2a	26	lufthansa	four seven	one		say	your	heading
	61	alitalia	six six	five		say	your	heading
	=======	=====	=======	=====	======	====	=====	=======
		call sign				"say"	"your"	"heading"

Pour les transcriptions 1a et 2a, on a regroupé les messages (nos. 1, 31, 67, 26 et 61) qui semblaient analogues. La dernière ligne du tableau montre les champs qui ont été identifiés comme composant le schéma générique "HDR". On constate qu'il y a d'abord un élément variable (indicatif d'appel), suivi d'un certain nombre de mots clés ("say","your", "heading").

On peut voir qu'il serait lourd de créer un tel schéma pour chacun des indicatifs d'appel susceptibles d'être appelés par le contrôleur. Cette remarque s'applique aussi à d'autres parties du message comme les caps, les altitudes, etc. On trouve alors avantageux de recourir à une propriété du mode de représentation des connaissances par schéma, c'est-à-dire la possibilité que l'une des cases d'un schéma soit une référence à un autre schéma. Pour l'exemple précédent, le seul élément dont la formulation peut varier est l'indicatif d'appel. On analyse la façon dont les indicatifs d'appel sont énoncés en utilisant la méthode illustrée au tableau 5 et on incorpore les résultats de cette analyse dans un schéma appelé CALLSIGNS (indicatifs d'appel). On obtient alors le schéma préliminaire complet montré à la figure qui suit.

Tableau VI Représentation du schéma HDR



On élabore les autres schémas de la même façon. On obtient alors un certain nombre de schémas correspondant à chacune des catégories de messages observés dans les transcriptions. On retrouve à l'annexe E un résumé de tous les schémas (et sous-schémas associés) pour les exercices simples.

6.4 Validation de l'utilisation des schémas dans le cas du même exercice exécuté par des élèves contrôleurs différents

On veut vérifier à quel point les schémas qui ont été élaborés décrivent les messages émis par les élèves contrôleurs. On veut surtout savoir si le modèle est stable d'un élève contrôleur à un autre. Le tableau suivant montre les schémas utilisés dans chacun des exercices ainsi que le % des messages schématisés (i.e. le % des messages totaux compris dans les deux exercices, à l'exclusion des messages engagés dans la gestion des erreurs) qu'ils représentent.

Tableau VII Proportion des messages décrits par les divers schémas

Nom du schéma	Transcription 1a	Transcription 2a	Cumulatif, en % du total
K	15,6 %	14,3 %	29,9 %
VSD	7,8 %	5,2 %	42,9 %
ALD/F	6,5 %	5,2 %	54,5 %
POD	5,2 %	6,5 %	66,2 %
HDR	6,5 %	2,6 %	75,3 %
VSD	6,5 %	2,6 %	84,4 %
ALD	3,9 %	3,9 %	92,2 %
HDD	1,3 %	2,6 %	96,1 %
ALR	-	1,3 %	97,4 %
OC	-	1,3 %	98,7 %
SB	-	1,3 %	100 %

On voit que les schémas proposés permettent de modéliser économiquement une part importante des messages provenant de chacun des deux élèves contrôleurs, i.e. qu'un nombre restreint de schémas décrit un grand nombre d'échanges. De plus, on voit que le modèle est stable entre les deux élèves contrôleurs : on peut modéliser leurs messages en utilisant à peu près les mêmes schémas.

6.5 Représentation d'un appareil

Pour compléter le modèle préliminaire du dialogue, on doit définir les objets sur lequel le dialogue va agir, i.e. la situation d'un appareil. On peut visualiser un appareil comme étant un objet qui se déplace suivant une trajectoire à travers un espace d'états (i.e. le micromonde de l'appareil), à partir de sa prise en charge par l'élève contrôleur jusqu'à son transfert à un autre centre de contrôle. Les états sont les paramètres dont tiennent compte les messages échangés entre le pseudo-pilote et l'élève contrôleur, ainsi que divers autres paramètres décrivant la situation d'un appareil donné. En fait, on propose ici que l'on peut considérer le dialogue entre le pseudo-pilote et l'élève contrôleur comme étant un outil permettant de faire évoluer l'appareil (ou l'objet) entre des états successifs. La figure suivante montre les divers états d'un appareil qui apparaît dans un des exercices, représenté sous forme d'objet.

Tableau VIII Représentation des divers états d'un appareil

Modele d'un appareil:

```
| Callsign: Lufthansa 471     |
| Previous                    |
| sector:   Westport          |
|                             |
|          Current  Next      |
|          =======  ====      |
| Sector  :Eastview Avalanche |
| Altitude:FL 280             |
| Heading :   355             |
| Position:15 m. S J565       |
| Altim.in.                   |
| Frequenc:   133.5   132.35  |
| Nor. nav:  N                |
|                             |
| Follow up: SIO Avalan.      |
```

On remarque que l'objet qui décrit l'appareil peut aussi être considéré comme un schéma. On peut donc facilement

intégrer cette représentation de l'appareil au modèle du langage. A l'intérieur du schéma, les différents champs constituent les états de l'appareil. Dans la plupart des cas, la valeur courante d'un état (notée "current" dans la figure) provient de l'instanciation d'un sous-schéma qui lui est associé. Le sous-schéma correspond à une partie d'un message qui a été reçu. La signification des différents états est :

- Callsign : l'indicatif d'appel de l'appareil, ici Lufthansa 471. Cet état est fixé pour toute la durée d'un exercice.

- Previous sector : secteur d'où provient l'appareil, ici Westport.

- Sector, altitude, heading, position, altimeter information, frequency : informations sur la situation courante ("current") ou prochaine ("next") de l'appareil.

- Normal navigation : l'étiquette est Y quand l'appareil assure sa propre navigation et N quand il est sous guidage par un contrôleur.

- Follow up list : pile des actions qui doivent être prises. Ici, l'appareil va entrer sous la juridiction d'un nouveau contrôleur (celui d'Avalanche). Il devra signaler sa présence, d'où le message SIO en attente d'être émis. Par ailleurs, quand un appareil reçoit une directive, il doit émettre un accusé de réception. L'accusé de réception reste sur la pile tant qu'il n'a pas été transmis.

6.6 Relation entre le dialogue et les états d'un appareil

Si on veut déterminer l'effet du dialogue sur l'évolution d'un appareil (ou de sa représentation dans l'espace d'états), on peut essayer de mettre en relation les schémas préliminaires du dialogue et la représentation de l'appareil. C'est ce qu'on fait à la figure suivante.

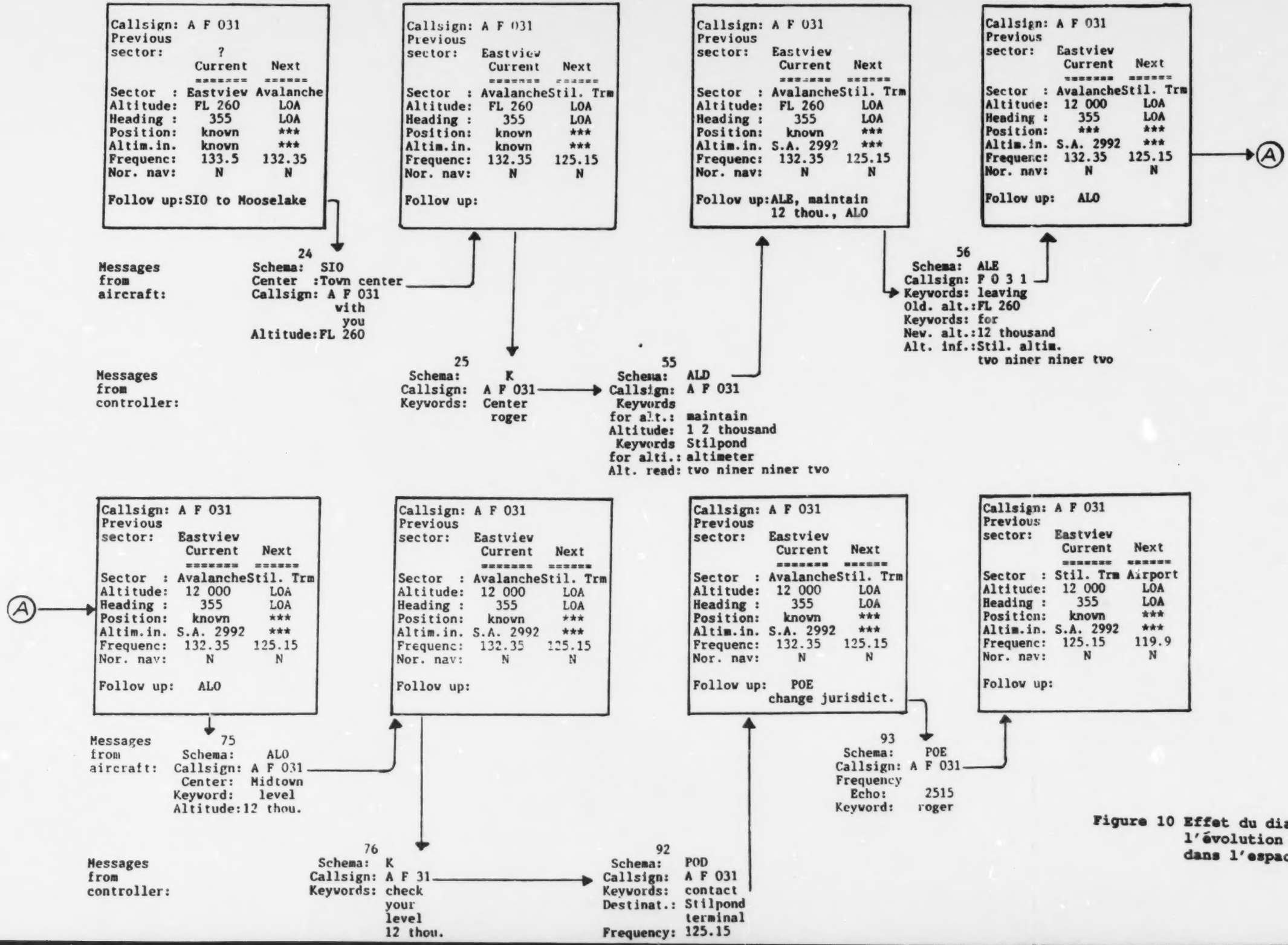


Figure 10 Effet du dialogue sur l'évolution d'un appareil dans l'espace d'états

Les encadrés représentent la "trajectoire" d'un appareil dans un espace d'états. Le trait continu représente la suite des états et des transitions entre états (occasionnés par les messages qui sont échangés). On a ainsi un premier aperçu du modèle du langage, où les relations entre les schémas représentant le langage opératif et le modèle d'un appareil apparaissent clairement.

Dans ce chapitre, on a élaboré une partie du modèle du langage pour les exercices simples. Ce modèle est incomplet : il manque les mécanismes d'instanciation des schémas, les mécanismes de gestion des erreurs ainsi que les règles d'évocation de schémas successifs. Le chapitre qui suit comble ces lacunes.

7. SECONDE ÉTAPE DU DÉVELOPPEMENT DU MODÈLE POUR LES EXERCICES SIMPLES

Ce chapitre permet de compléter le modèle de langage pour les exercices simples qui a été ébauché au chapitre précédent. On y parvient en identifiant les mécanismes d'instanciation des schémas, de gestion des erreurs de dialogue et d'évocation de schémas successifs.

7.1 Mécanismes d'instanciation des schémas

On veut dégager des règles qui permettront d'aider à instancier certains des champs des schémas. On identifie d'abord des sources additionnelles d'information disponible lors de l'exécution de la tâche et comment on peut en tirer parti. On examine ensuite l'instanciation de divers champs numériques.

7.1.1 Identification des sources d'information

On peut tirer avantage du fait que la tâche est constituée en réalité d'un scénario et que l'élève contrôleur évolue à l'intérieur d'un univers bien défini (i.e. le micromonde du contrôleur). De plus, comme à tout moment l'élève contrôleur peut observer un appareil sur son écran radar, l'ordinateur du simulateur de trafic aérien possède aussi des informations sur les appareils qui évoluent dans l'espace aérien.

L'exemple de la section 6.6 montre que seuls quelques uns des états de l'appareil (par exemple, l'altitude) sont influencés par un message donné. De façon plus générale, le tableau suivant montre les relations entre chacun des schémas obtenus de l'analyse sommaire et les états de l'appareil. On distingue plusieurs cas. D'abord, un élément commun entre un schéma (ou message) donné et un des états de l'appareil peut être simplement mentionné et n'entraîne pas de changement dans le micromonde de l'appareil. Cette relation est identifiée par "A" dans le tableau. Un message reçu peut faire référence à l'indicatif d'appel de l'appareil, mais comme cet état est

constant, il n'y a pas de changement dans le micromonde de l'appareil. Ensuite, on a le cas où un élément commun entre le schéma et le micromonde de l'appareil change le micromonde de l'appareil. Il y a deux possibilités. D'abord, le changement est entièrement prévisible à priori (par exemple à cause du scénario de l'exercice) et l'évocation du schéma ne fait "qu'entériner" le changement. C'est le cas notamment pour un transfert à un autre centre de contrôle. Le scénario de l'exercice prévoit ce changement et tous les détails sont connus soit à l'avance (comme la fréquence du nouveau centre de contrôle), soit dynamiquement par l'ordinateur de la simulation (comme la position de l'appareil). Ces éléments sont identifiés par "B" dans le tableau. Enfin, le changement peut être imprévisible à priori. C'est le cas d'un changement d'altitude, par exemple. Le message doit donc fournir toute l'information requise. Ces éléments sont identifiés par "C" dans le tableau. L'attribution des cotes A, B ou C est justifiée à l'annexe F.

Tableau IX **Relations entre la représentation d'un appareil et les schémas**

Etats de l'appareil	Schémas HDR	HDD	VRD	VSD	ALDF	ALD	ALR	POD	K	OC	SB
Callsign	A	A	A	A	A	A	A	A	A	A	A
Previous sector											
Sector: current											
next								B			
Altitude: current											
next					C	C					
Heading: current											
next		C	C	C							
Position: current			B								
next											
Altimeter informat.: current											
next					B	B					
Frequency								B			
Norm. nav			B								
Follow-up	HDA	HDE	VRE	VSE	ALEF	ALE	ALA	POE			

L'existence de sources alternatives d'information donne la possibilité de compenser les erreurs introduites par le reconnaisseur de parole. Il faut toutefois user de prudence

avec ce genre de méthode : en effet, une utilisation abusive de ce procédé pourrait permettre de corriger les erreurs de l'élève contrôleur.

7.1.2 Règles pour l'instanciation de certains sous-schémas

Au chapitre précédent, on a présenté divers sous-schémas contenant des champs numériques ("altitude", "heading", etc.). On peut utiliser les connaissances sur la situation de contrôle aérien, l'état d'un appareil donné, l'historique du dialogue, etc. pour créer des règles afin de limiter les valeurs que peuvent prendre ces champs. Le tableau suivant identifie quelques unes de ces règles.

Tableau X **Limites pour l'instanciation de divers sous-schémas**

Sous-schéma	Règle et justification
Next altitude	**minimum altitude (current sector) < next altitude = maximum altitude (current sector)** Justification: Un contrôleur ne peut assigner à un appareil une altitude qui n'est pas comprise dans le secteur qu'il contrôle. Note: il existe des altitudes préférentielles pour effectuer un transfert à un autre secteur. Un modèle plus élaboré pourrait en tenir compte.
Altitude to be reported	**si next altitude > current altitude, alors** **current altitude > altitude to be reported = next altitude** **si next altitude < current altitude, alors** **current altitude < altitude to be reported = next altitude** Justification: l'item "altitude to be reported" marque une étape lors du changement vers l'item "next altitude".
Heading to reintercept	Règle: **Le cap doit permettre d'intercepter l'objectif décrit dans le message qui a été communiqué à l'appareil.**

Ce tableau est incomplet. On pourrait vraisemblablement dégager

d'autres règles pour d'autres champs (numériques ou non) en utilisant les connaissances d'un contrôleur expérimenté.

7.2 Gestion des erreurs de dialogue

Cette section examine comment sont gérées les erreurs dans les exercices. Pour enrichir les données, on considère en plus des exercices de contrôle en basse altitude les exercices de contrôle "de terminal". On identifie alors les processus de gestion des erreur qui sont utilisés. On veut dégager des règles susceptibles d'être utilisées lors du dialogue entre l'élève contrôleur et le pseudo-pilote. On entend par erreurs les événements se produisant dans le dialogue et qui sont susceptibles d'aiguiller l'un ou l'autre des interlocuteurs dans une mauvaise direction. Il s'agit d'événements "graves" (par exemple le nouveau cap est mal compris), par rapport à d'autres erreurs qui sont susceptibles d'être tolérées par le modèle basé sur des schémas (par exemple, une abréviation de l'indicatif d'appel d'un appareil).

7.2.1 Classification des situations d'erreur

L'examen des transcriptions montre que l'on retrouve diverses situations où des erreurs de dialogue se produisent. On propose dans le tableau qui suit une classification de ces erreurs de dialogue, en ordre approximatif de complexité. On trouve dans le tableau toutes les situations où des erreurs de dialogue ont été observées dans les exercices simples (contrôle en altitude peu élevée) et dans les exercices complexes (contrôle de terminal).

Tableau XI Classification des erreurs de dialogue

Type d'erreur	Messages relatifs à une erreur et provenant de l'élève-contrôleur	Messages relatifs à une erreur et provenant du pseudo-pilote	Total
Oubli et erreur du pilote	1 (0,5 %)	1 (0,4 %)	2 (0,4 %)
Erreur avec autocorrection à l'intérieur du même message	12 (5,5 %)	-	12 (2,7 %)
Erreur avec autocorrection lors de messages successifs	1 (0,5 %)	1 (0,4 %)	2 (0,4 %)
Demande de répétition	3 (1,4 %)	3 (1,3 %)	6 (1,3 %)
Identification et résolution d'incertitude	1 (0,5 %)	6 (2,6 %)	7 (1,6 %)
Total	18 (8,3 %)	11 (4,8 %)	29 (6,5 %)

On peut décrire les situations d'erreurs de dialogue comme :

. Oublis et erreurs : ces erreurs se produisent quand un interlocuteur reçoit les bonnes instructions mais qu'il fait une erreur "de manipulation" (par exemple le pilote suit le mauvais cap).

- Erreur avec autocorrection à l'intérieur du même message : ces erreurs se produisent quand un interlocuteur fait une erreur, s'en rend compte, et se corrige à l'intérieur du même message.

- Erreur avec autocorrection lors de messages successifs : ces erreurs se produisent quand un interlocuteur fait une erreur dans un message donné, s'en rend compte après que le message soit terminé, et envoie un autre message pour corriger l'erreur.

- Demandes de répétition : ces erreurs se produisent quand un des interlocuteurs demande à l'autre de répéter une information.

- Identifications et des résolutions d'incertitudes : ces erreurs se produisent quand un des interlocuteurs identifie une incertitude et entreprend une démarche pour la résoudre. Une incertitude n'est pas nécessairement une conséquence immédiate de l'échange de messages entre l'élève contrôleur et le pilote.

7.2.2 Processus de gestion des erreurs

Cette section examine comment s'effectue la gestion du dialogue pour identifier et corriger certains types d'erreurs. Comme le but que l'on recherche est de construire un modèle du dialogue qui soit souhaitable pour un pseudo-pilote automatique, on choisit d'éliminer le cas où le pseudo-pilote produit lui-même une erreur (cas des oublis et des erreurs du pilote). On élimine aussi les cas où le pseudo-pilote effectue des autocorrections sur les messages qu'il émet (la machine ne doit pas faire d'erreurs).

On analyse les processus de gestion des erreurs pour les cas restants de la même façon que l'on a analysé la constitution des schémas. On a d'abord regroupé tous les exemples de gestion des erreurs du même type; ensuite, on a essayé d'en extraire un ou plusieurs patrons. Le détail de ces analyses se trouve à l'annexe G. Les tableaux qui suivent montrent les mécanismes qui ont été observés pour chacun des cas.

Tableau XII Méthodes de gestion des erreurs avec autocorrection

Catégorie d'erreur	Mécanismes de gestion des erreurs qui ont été observés		
Autocorrection à l'intérieur du même message (par l'élève-contrôleur)	BLOC SEMANTIQUE ERRONE	+ MOT CLE : CORRECTION	+ BLOC SEMANTIQUE CORRIGE
	BLOC SEMANTIQUE INCOMPLET	+ MOT CLE: CORRECTION	+ BLOC SEMANTIQUE CORRIGE
	ELEMENT OU MOT INCOMPLET	+ MOT CLE: CORRECTION	+ BLOC SEMANTIQUE CORRIGE
Autocorrection lors de messages successifs (par l'élève contrôleur)	MESSAGE INITIAL ERRONE	+ MESSAGE SUCCESSIF CORRIGE CONTENANT LE MOT-CLE "CORRECTION"	

A partir du tableau précédent, on dégage les règles requises pour la gestion d'une erreur avec autocorrection à l'intérieur du même message. Ces règles sont :

- Règle : si on rencontre le mot CORRECTION dans un message, le bloc sémantique suivant ce mot remplace celui de même nature qui précède le mot CORRECTION.

. Règle : si deux blocs sémantiques de même nature se suivent dans un même message, le dernier bloc sémantique sera celui considéré valide.

Pour les erreurs avec autocorrection lors de messages successifs, il faut prévoir deux cas. Soit le pseudo-pilote automatique aurait relevé l'erreur (à la différence du pseudo-pilote humain qui peut occasionnellement ne pas détecter l'erreur avant que l'élève contrôleur ne la relève) et un mécanisme d'identification et de résolution d'incertitude aurait été invoqué. Soit le pseudo-pilote n'aurait pu relever l'erreur et dans ce cas, le contenu du message successif contenant la correction remplace simplement celui du message erroné. On doit donc ajouter une seule règle pour traiter ce dernier cas :

. Règle : si un message contient le mot-clé CORRECTION et qu'il ne s'agit pas d'une autocorrection du même message, alors ce message est traité comme une nouvelle instruction pour l'appareil.

Le prochain tableau traite le cas des demandes de répétitions provenant de l'élève contrôleur.

Tableau XIII Méthode de gestion des demandes de répétition originant de l'élève-contrôleur

Catégorie d'erreur	Mécanismes de gestion des erreurs qui ont été observés
Demandes de répétition (par l'élève-contro-leur	DESTINATAIRE + MOTS-CLES: "SAY AGAIN"

On peut dégager la règle suivante pour gérer les demandes de répétition de l'élève contrôleur :

. Règle : si le message "Destinataire SAY AGAIN" est

émis par l'élève contrôleur, le dernier message émanant d'un appareil doit être répété.

On passe maintenant aux demandes de répétition provenant du pseudo-pilote. Le tableau suivant montre les résultats de l'analyse des transcriptions.

Tableau XIV Méthode de gestion des demandes de répétition originant du pseudo-pilote

Catégorie d'erreur	Mécanismes de gestion des erreurs qui ont été observés
Demandes de répétition (par le pseudo-pilote)	Echo par le pseudo-pilote du message: Callsign + Bloc(s) sémantique(s) valide(s) (facultatifs) + MOTS-CLES "SAY AGAIN" + Bloc sémantique douteux

On modélise cette situation avec la règle suivante :

- Règle : si la valeur d'un bloc sémantique est douteuse ou incomprise, le pseudo-pilote peut émettre une demande de répétition pour ce bloc sémantique.

Il reste à considérer le cas où la gestion d'une erreur est la plus complexe, i.e. quand l'un des interlocuteurs identifie une ambiguité et implique l'autre individu (i.e. le correspondant) pour lever l'ambiguité. Le tableau qui suit montre deux types de mécanismes qui ont été observés.

Tableau XV Mécanismes d'identification et de résolution d'ambiguité

Dialogue	Commentaires
MESSAGE EMIS	Un interlocuteur émet un message quelconque.
BLOC SEMANTIQUE+ "CONFIRM"	Le correspondant répond en identifiant un élément douteux et demande confirmation.
"CONFIRM" OU "NEGATIVE" OU "CORRECTION" + BLOC SEMANTIQUE	L'interlocuteur confirme ou infirme l'élément douteux. La gestion de l'erreur est alors complétée.
MESSAGE EMIS	Un interlocuteur émet un message quelconque.
BLOC SEMANTIQUE+ "CONFIRM"	Le correspondant répond en identifiant un élément douteux et demande confirmation.
L'interlocuteur ne répond pas.	
DEMANDE (LANGAGE NATUREL)	Après un certain temps, l'interlocuteur demande au correspondant s'il a effectué la manoeuvre demandée.
REPONSE (LANGAGE NATUREL)	Le correspondant informe l'interlocuteur qu'il a demandé confirmation.
BLOC SEMANTIQUE+ "CONFIRM"	L'interlocuteur confirme ou non l'élément. La gestion de l'erreur est alors complétée.

On voit que pour le premier cas, le processus de gestion d'erreur est simple. On peut facilement modéliser ce mécanisme avec la règle suivante :

- Règle : si la valeur d'un bloc sémantique est douteuse, l'interlocuteur peut formuler une demande de confirmation du bloc sémantique en question.

On note aussi qu'en dépit de la diversité des formulations de la part de l'interlocuteur, la règle permettant de formaliser ce mécanisme est quasi identique à celle traitant des demandes de répétition provenant du pseudo-pilote.

Le second mécanisme reflète une anomalie : une incertitude a été identifiée et le correspondant (ici le pseudo-pilote) a tenté de la résoudre. La transcription montre qu'à ce moment, l'interlocuteur (ici l'élève contrôleur) a entrepris une conversation avec un autre interlocuteur (peu importe qu'il s'agisse d'un autre appareil ou d'un centre de contrôle). La demande de résolution d'incertitude est demeurée en suspend, et le pseudo-pilote n'a pas réitéré sa demande et s'est campé sur ses positions. Eventuellement, l'élève contrôleur s'est enquis des intentions du pseudo-pilote (l'élève contrôleur soit n'avait pas entendu la requête initiale du pseudo-pilote, soit l'avait mise en attente et l'avait oubliée). Les deux interlocuteurs ont dû recourir à des échanges en langage naturel pour reprendre le dialogue là où il avait été interrompu.

Même si cette situation semble plus complexe, on peut probablement éviter les complications amenées par la reprise hésitante du dialogue entre l'élève contrôleur et le pseudo-pilote. En effet, si le pseudo-pilote avait relancé l'élève contrôleur après que celui-ci ait terminé la conversation entreprise avec un autre interlocuteur, on peut croire que le gestion de l'erreur se serait effectué comme dans le premier cas. On prévoit donc la procédure suivante pour permettre la gestion de l'erreur de façon à le garder dans un cadre bien défini :

- Tant que l'incertitude sur la valeur d'un bloc sémantique persiste, on doit :
  - Attendre que l'échange en cours entre l'élève contrôleur et un autre interlocuteur soit terminé.

. Utiliser la règle précédente (celle de la méthode simple d'identification et du résolution d'incertitude) pour résoudre la situation.

Même si ce processus de gestion d'erreur est très primitif, il force les interlocuteurs à résoudre l'incertitude pour permettre la poursuite de la tâche du pseudo-pilote. Cet entêtement du processus de gestion du dialogue aurait peut-être pu permettre de résoudre un incident où un élève contrôleur persistait à appeler son interlocuteur Nordair alors que celui-ci avait mentionné à plusieurs reprises que son identification était Norontair.

7.3 Évocation de schémas successifs

Cette section examine comment on peut restreindre le choix des schémas parmi lesquels doit être classifié un message donné. On vise ainsi à améliorer la performance du système de reconnaissance de la parole.

7.3.1 Choix du destinataire d'un message

On peut restreindre le nombre de candidats auxquels peut être destiné un message émis. En effet, seuls les appareils qui sont dans la zone contrôlée sont susceptibles d'être appelés et de répondre. Si par erreur un message était envoyé à un appareil qui n'est pas dans la zone contrôlée (et qu'il ne s'apprête pas à y pénétrer), le message devrait rester sans réponse : en effet, l'appareil n'est pas sur la fréquence radio de l'élève contrôleur. On obtient donc la règle suivante :

. Règle : seuls les appareils présents dans la zone contrôlée (ou s'apprêtant à y pénétrer) sont susceptibles de recevoir un message de l'élève contrôleur.

7.3.2 Succession des schémas lors de l'évolution d'un appareil

Quand un appareil évolue dans un secteur contrôlé, sa trajectoire et les messages qu'il est susceptible de recevoir

dépendent de plusieurs facteurs. Par exemple, s'il évolue dans une zone de basse altitude (ici, Avalanche), les messages relatifs à l'approche finale sont exclus de son micromonde. De plus, les messages permissibles dépendent aussi de sa position à l'intérieur de la zone contrôlée. Ainsi, un appareil qui arrive tout juste d'une autre zone et qui déclare sa position au moyen d'un message approprié n'est pas susceptible de recevoir un message de transfert à une autre zone avant un certain temps. Enfin et surtout, il semble inévitable que les catégories de messages qu'un appareil reçoit dépendent du scénario de l'appareil et de l'exercice dans lequel il est impliqué. On pose donc les hypothèses suivantes :

. L'ordre d'évocation des schémas successifs dépend en partie des schémas eux-mêmes : certains schémas ne peuvent simplement pas se trouver dans une partie précise d'un script, alors que d'autres sont presqu'inévitables.

. Des élèves contrôleurs différents vont utiliser sensiblement les mêmes schémas pour contrôler des appareils qui sont dans des situations analogues.

Pour vérifier ces hypothèses, on reprend les transcriptions et on construit le tableau suivant :

Tableau XVI Evocation de schémas successifs

Séquence des messages pour les appareils suivants:									
N900PB		AC 951		LUF 471		AF 031		CAN 63	
Transcr.: 1a	2a	Transcr.: 1a	2a	Transcr.: 1a	2a	Transcr.: 1a	2a	Transcr: 1a	2a
				SIO K	SIO	SIO K	SIO K	SIO K	SIO SB
HDR				HDR	HDR	ALD	ALD		
HDA				HDA	HDA	ALE	ALE	( )	
VSD				VSD	HDD	ALO	K		ALR
( )				( )	HDE	K			ALA
ALDF	ALDF	ALD	ALD	ALDF	ALDF				K
ALEF	ALEF	ALE	ALE	ALEF	ALEF			ALD	ALD
ALF	ALF			ALF				ALE	ALE
	K			K					
VRD	VRD			VRD	VRD				
VRE	VRE			VRE	VRE				
[VRD]				ALO	ALF				
K				( )					
				ALO					
				K					
POD	POD	POD	POD	POD	POD	POD	POD		POD
POE	POE	POE	POE	POE	POE	POE	POE		POE
			K						

Le tableau montre d'abord, pour un certain nombre d'appareils, les schémas qui ont été échangés entre l'élève contrôleur et le pilote. On a inclus les schémas modélisant les réponses du pilote, soient SIO, HDA, ALEF, ALF, ALO et VRE, bien que l'on ne considérera pas ces schémas dans cette analyse. Enfin, on constate qu'il semble manquer une partie des séquences de schémas pour certains appareils : cela s'explique par le fait que les enregistrements des différents exercices ont débuté

à des moments différents, alors que l'exercice était déjà en cours.

La première hypothèse semble vérifiée : en effet, certains schémas semblent se retrouver exclusivement au début ou à la fin du dialogue entre chaque élève contrôleur et chaque pilote (par exemple, SIO, POD, POE). C'est même une conclusion triviale à cause de la nature même de ces schémas.

Il existe aussi une corrélation très forte entre les schémas évoqués dans le cas du même appareil contrôlé par deux élèves contrôleurs différents. Ils utilisent sensiblement les mêmes schémas, dans un ordre à peu près semblable. Cela tend aussi à confirmer la seconde hypothèse.

L'examen des transcriptions confirme aussi que tous les appareils n'ont pas besoin de tous les schémas disponibles. Par exemple, AF 031 et CAN 63 n'utilisent qu'une partie des schémas utilisés par LUF 471 et N900PB. Par ailleurs, il semble évident qu'à moins d'une erreur toujours possible de la part d'un élève contrôleur, seul le sous-ensemble pertinent de tous les schémas disponibles sera utilisé pour effectuer le contrôle de l'appareil. On peut donc dire qu'il existe un ensemble de schémas disponibles pour accomplir une tâche de contrôle donnée : cet ensemble est constitué de "sous-séquences" de schémas nécessaires pour réaliser la tâche. Par exemple, le transfert d'un appareil à une autre zone nécessite la séquence POD-POE, suivie accessoirement d'un schéma K.

On propose ici des règles particulières concernant certains schémas de même que des règles plus générales concernant l'utilisation d'un ensemble de schémas pour un appareil particulier :

- Règle : lorsqu'un appareil arrive dans une zone de contrôle, il envoie généralement un message SIO. La réponse de l'élève contrôleur peut être un message K, SB ou un autre message.

- Règle : si un appareil est à une altitude différente de l'altitude préférentielle d'entrée du prochain secteur (i.e. sa destination), une sous-séquence de changement d'altitude devra être utilisée en temps voulu et avant un message POD.

- Règle : si un appareil suit un cap différent du cap

d'entrée préférentiel du prochain secteur, une sous-séquence de changement de cap devra être utilisée en temps voulu et avant un message POD.

. Règle : la dernière sous-séquence dans le contrôle d'un appareil est une sous-séquence comprenant POD-POE, suivie accessoirement d'un message K.

7.4 Modèle résultant

Cette section présente une partie du modèle de langage qui a été développé jusqu'ici. On a maintenant :

. Un squelette composé de blocs qui illustre l'architecture d'un système de gestion de dialogue et qui incorpore une partie des mécanismes de gestion d'erreurs.

. Des schémas qui intègrent les autres mécanismes de gestion d'erreurs et les diverses autres modifications notées jusqu'ici.

Il convient de souligner ici que le modèle proposé n'est qu'une des conceptualisations possibles de l'architecture du modèle de dialogue. En effet, pour incorporer les améliorations qui ont été identifiées dans ce chapitre, il a fallu fixer un certain nombre de choses qui étaient restées jusqu'ici à l'état de concept. On a donc matérialisé ces concepts sous une forme donnée : on ne prétend pas ici que cette forme soit unique, ni quelle soit la meilleure. Simplement, elle permet de rassembler les éléments obtenus en un ensemble cohérent susceptible d'être utilisable. La figure qui suit illustre d'abord le squelette composé de blocs qui constitue l'architecture du système de gestion du dialogue.

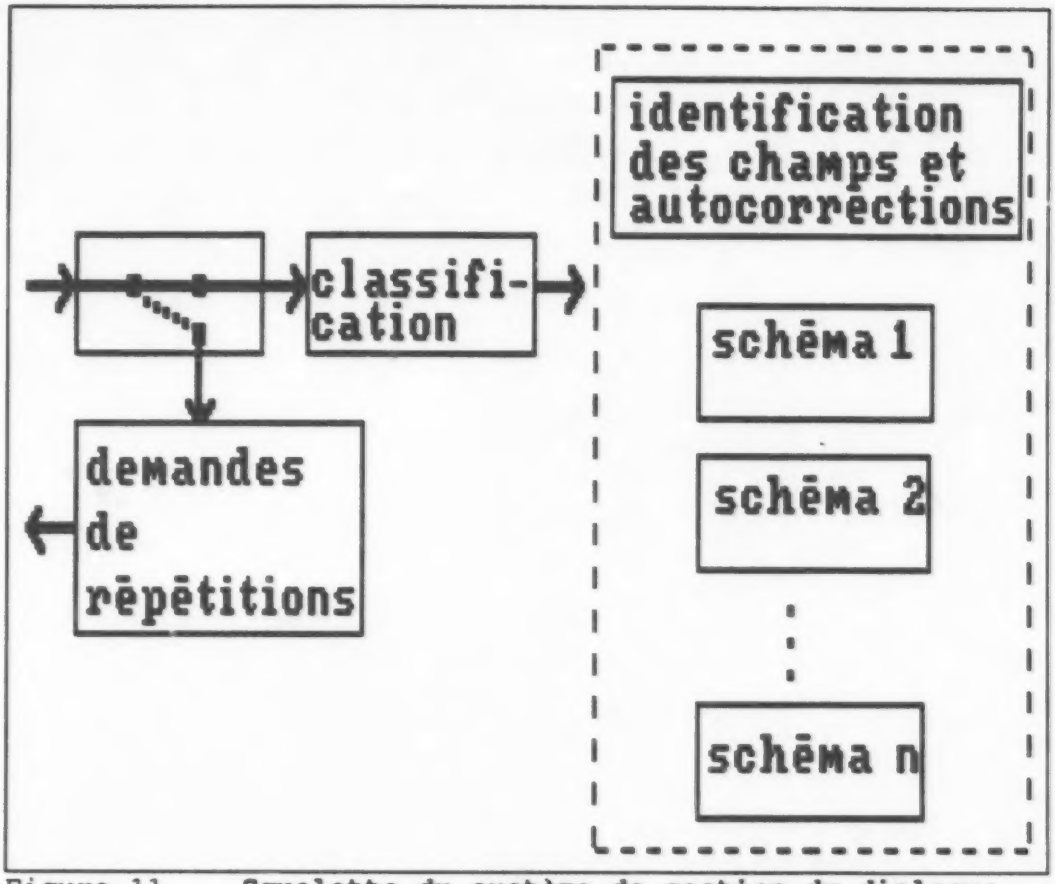


Figure 11 **Squelette du système de gestion du dialogue**

On remarque que les règles qui permettent la gestion des erreurs ont été distribuées dans le système. Cette distribution correspond fonctionnellement aux mécanismes observés :

- Les demandes de répétition de l'élève contrôleur sont traitées sans même que l'on tente d'identifier ce dont il est question.
- Les autocorrections semblent être traitées lors de l'instanciation du schéma qui a été identifié.
- L'identification et la résolution d'incertitude sont traitées au niveau des champs concernés pour le schéma auquel le message appartient.

La figure suivante montre un exemple de schéma qui incorpore certaines des améliorations qui ont été identifiées dans ce chapitre :

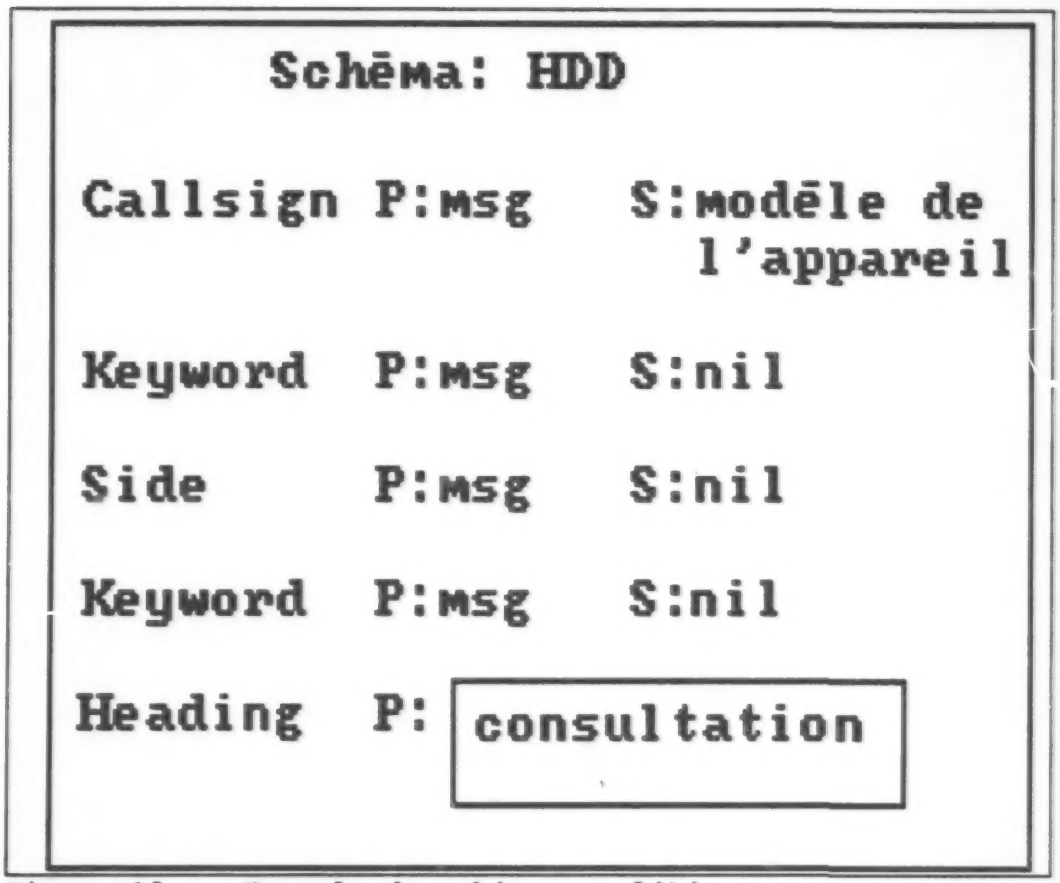


Figure 12 **Exemple de schéma complété**

Pour chacune des cases du schéma, la lettre P désigne la source principale pour l'information, et la lettre S désigne une source secondaire. Par exemple, pour le champ Callsign, si la portion correspondante du message est déficiente, on peut utiliser le modèle de l'appareil à qui le message est destiné pour effectuer une correction. Les champs qui suivent (Keyword, Side, Keyword) contiennent des mots-clés ou des valeurs qui n'ont pas un contenu sémantique essentiel, pour l'instant. Le dernier champ (Heading) illustre le dernier mécanisme de gestion d'erreur : l'identification et la résolution d'incertitude. L'examen des transcriptions montre que pour qu'une incertitude

soit identifiée, l'incertitude doit être significative et le pseudo-pilote doit être en mesure de la relever. Par exemple, une erreur concernant une des lettres de l'indicatif d'appel (Callsign) ne justifie pas, en général, le déclenchement du processus. Par contre, une incertitude concernant des informations vitales est relevée et doit être résolue. L'examen des transcriptions montre que ces informations vitales sont, dans le cadre des exercices qui ont été examinés, l'altitude et la direction. Malheureusement, la seule source pour ces informations est constituée par le message lui-même. Ce dernier peut soit être correct, soit être erroné à cause d'une erreur du reconnaisseur de parole, soit être erroné à cause d'une erreur de l'élève contrôleur. On peut donc utiliser certaines des règles proposées pour l'instanciation de certains sous-schémas pour valider, et éventuellement corriger, ces champs. C'est la tâche du processus "Consultation" dans l'exemple précédent. Ce processus reçoit la valeur qui est extraite du message pour le champ concerné et vérifie si cette valeur est plausible compte tenu des règles d'instanciation pour ce sous-schéma (altitudes et routes préférentielles, probabilité du message par rapport à la situation courante de l'appareil et à l'historique du dialogue, etc.). Le processus Consultation peut être considéré, par exemple, comme un système expert opérant en mode de chaînage arrière : à partir d'une conclusion (i.e. à partir d'une altitude ou d'un cap proposé), on tente de déterminer si les conditions menant à cette conclusion sont vérifiées. Le cas échéant, le processus Consultation indique que la conclusion proposée est vraisemblable. Dans le cas contraire, le processus indique que la conclusion n'est pas plausible et peut éventuellement proposer une conclusion différente, plausible dans les circonstances.

Le modèle qui a été présenté ici décrit le langage utilisé lors d'exercices simples. Le chapitre qui suit examine si on peut utiliser ce modèle dans une situation plus complexe, i.e. le contrôle de terminal. On aura ainsi fait un tour d'horizon sur la modélisation du langage utilisé par les élèves contrôleurs.

8. EXTENSION DU MODÈLE RÉSULTANT

Dans les deux chapitres précédents, on a élaboré un modèle de dialogue qui permet de comprendre les échanges entre un élève contrôleur et un pilote dans un exercice simple.

On veut maintenant vérifier si ce modèle simple est extensible, i.e. si on peut l'adapter à une situation plus complexe. Comme on l'a mentionné au chapitre 3, la situation plus complexe est celle du contrôle de terminal. On a donc réalisé deux enregistrements d'exercice de contrôle de terminal à l'Institut de formation de Transports Canada à Cornwall. Chaque enregistrement dure 45 minutes et correspond à l'accomplissement d'un exercice par un élève contrôleur différent. Les transcriptions orthographiques sont appelées 3a et 4a et se trouvent à l'annexe H.

On veut aussi voir comment on peut relâcher les contraintes qu'on s'était fixées sur la concaténation des messages à l'intérieur d'une transmission, et sur le nombre de destinataires qu'on peut retrouver dans une transmission.

8.1 Adaptabilité du modèle

Pour vérifier si on peut adapter le modèle simple au contrôle de terminal, on doit :

- Vérifier l'ordre de grandeur et la composition du vocabulaire opératif pour le contrôle de terminal.
- Vérifier si les schémas du modèle simple peuvent être utilisés pour analyser une partie des échanges entre l'élève contrôleur et le pseudo-pilote, dans le contrôle de terminal.
- Vérifier si on peut adopter la même méthode d'analyse que dans le cas du modèle simple pour ajouter les schémas nécessaires pour compléter le modèle.

8.1.1 Examen de la longueur des transmissions en contrôle de terminal

Comme on l'a fait pour les exercices simples, on veut évaluer la longueur des transmissions émanant de l'élève contrôleur. La figure qui suit montre d'abord la distribution de la longueur des transmissions dans le cas du contrôle de terminal :

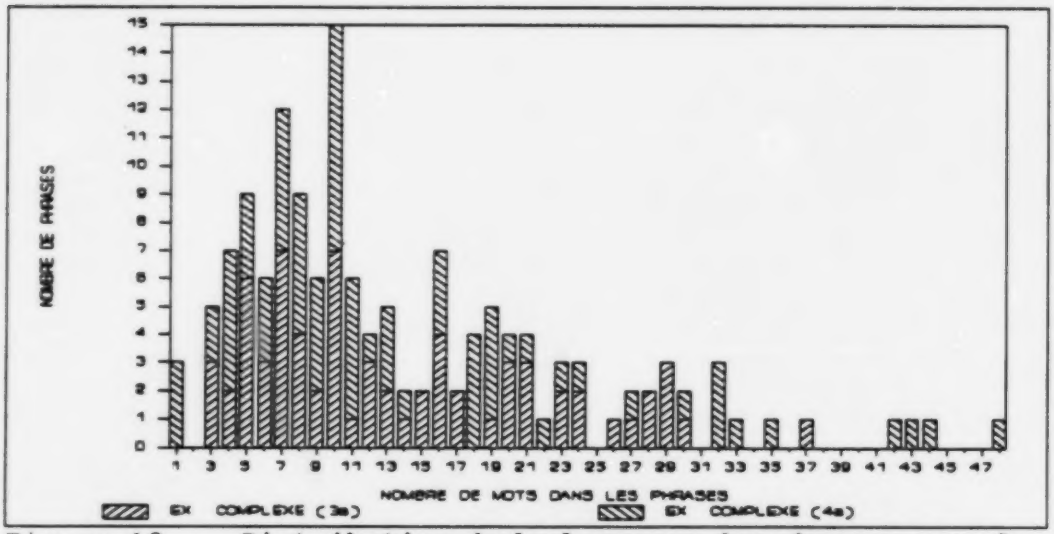


Figure 13 **Distribution de la longueur des phrases pour les exercices complexes**

On veut maintenant savoir comment la distribution de la longueur des phrases pour les exercices complexes se compare à celle des exercices simples. C'est ce qu'on fait à la figure qui suit.

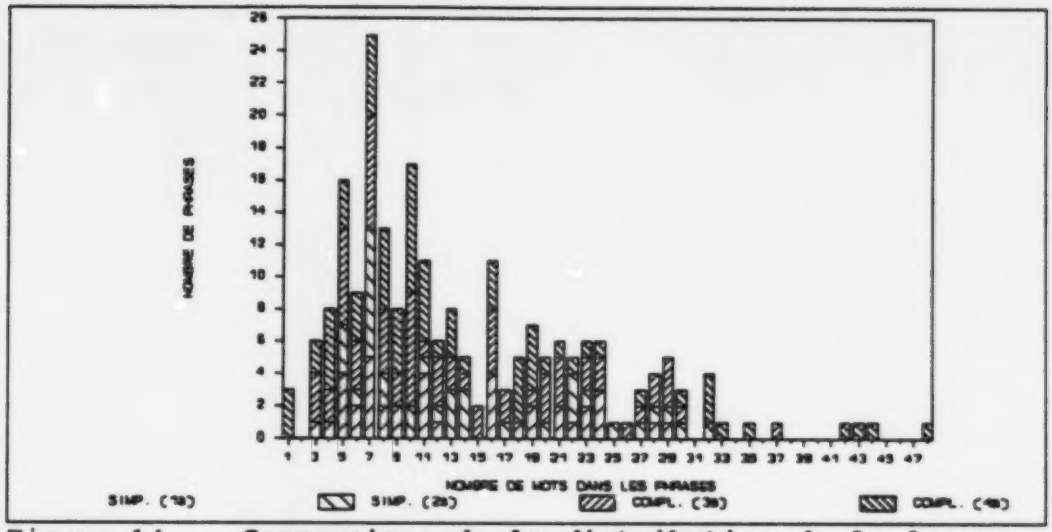


Figure 14 **Comparaison de la distribution de la longueur des phrases pour les exercices simples et les exercices complexes**

On voit que dans le cas des exercices simples et complexes qui ont été examinés, la majorité des transmissions ont entre 3 et 30 éléments de longueur.

8.1.2 Vocabulaire opératif pour le contrôle de terminal

Dans cette section, on évalue la taille du vocabulaire requis pour effectuer la tâche de contrôle de terminal. On procède de la même façon que pour le cas du modèle pour une tâche simple : on trouve le vocabulaire net opératif (incomplet) pour chacune des transcriptions. On trouve ensuite le vocabulaire net opératif (incomplet) combiné pour les deux transcriptions. Enfin, on ajoute un nombre standard de mots pour tenir compte des nombres, des centre de contrôle, des lignes aériennes, etc. Le tableau qui suit donne la taille du vocabulaire opératif net incomplet (la liste des mots du vocabulaire se trouve à l'annexe I) pour ces exercices.

Tableau XVII Taille du vocabulaire opératif net (incomplet) pour les exercices de terminal

	Transcription 3a	Transcription 4a
Vocabulaire opératif net (déduction des références aux nombres, lignes aériennes, etc.	60 mots	71 mots

On constate d'abord que les vocabulaires de chaque transcription sont à peu près du même ordre. Comme pour le modèle simple, on détermine quelle est la taille du vocabulaire requis pour l'ensemble des deux transcriptions et on obtient un vocabulaire combiné net de 81 mots. Il semble donc que des élèves contrôleurs différents utilisent sensiblement le même vocabulaire. On ajoute ensuite le nombre standard de mots requis (nombres, centres de contrôle, etc.) et on obtient :

Tableau XVIII Taille du vocabulaire opératif pour le contrôle de terminal

Taille du vocabulaire opératif net (incomplet)	Nombre d'éléments standards	Taille totale du vocabulaire opératif
81 mots	79 mots	160 mots

Si on prend une marge de sécurité de 25 %, on obtient un vocabulaire opératif total de (160 * 1.25) 200 mots pour le contrôle de terminal.

On fait une analyse supplémentaire pour déterminer la similarité des vocabulaires entre la tâche simple et la tâche complexe. On ne considère que le vocabulaire opératif net car le nombre d'éléments standards est le même dans les deux cas.

On constate que le vocabulaire opératif net incomplet permettant de décrire les deux tâches simultanément est de 85 mots, i.e. à l'exception de 4 mots le vocabulaire de la tâche simple est entièrement inclus dans celui de la tâche complexe. Ce résultat est intéressant car il donne les choix suivants :

- Utiliser un vocabulaire pour chaque type de tâche : on peut alors utiliser un vocabulaire de plus petite taille pour les tâches simples, ce qui peut favoriser une meilleure performance du reconnaisseur de parole.

- Utiliser un seul vocabulaire recouvrant la tâche simple et la tâche complexe : on diminue la complexité du système.

8.1.3 Vérification de l'utilité des schémas du modèle simple pour le contrôle de terminal

On peut évaluer l'utilité des schémas du modèle simple comme constituants du modèle du dialogue pour le contrôle de terminal en examinant quelle proportion des messages entre l'élève contrôleur et le pilote peuvent être "expliqués" ou analysés avec les schémas du modèle simple. Si on trouve que beaucoup de schémas sont "réutilisables" dans le contexte du contrôle de terminal, cela suggère que le modèle simple est robuste et qu'on peut peut-être lui faire prendre de l'expansion. Si par contre on trouve que les schémas du modèle simple sont à peu près absents du dialogue de contrôle de terminal, cela suggère qu'il faut un modèle différent pour chaque secteur de contrôle. Ce dernier cas rendrait problématique l'utilisation pratique du modèle.

On commence par extraire des transcriptions de contrôle de terminal les messages qui proviennent de l'élève contrôleur. On tente alors d'utiliser les schémas définis dans le modèle préliminaire pour décrire ces messages. Le tableau qui suit présente les résultats de cette analyse :

Tableau XIX Vérification de l'utilité des schémas du modèle simple pour le contrôle de terminal

Situation	Transcription 3a	Transcription 4a
Cas où les schémas du modèle simple suffisent à analyser les messages entre l'élève-contrôleur et le pseudo-pilote	55.7 %	59.4 %
Cas où les schémas du modèle simple doivent être complétés pour permettre d'analyser les messages	8.6 %	10.8 %
Cas où des schémas supplémentaires sont nécessaires pour permettre d'analyser les échanges entre l'élève-contrôleur et le pseudo-pilote.	35.7 %	29.7

On constate que près des deux tiers des échanges entre l'élève contrôleur et le pseudo-pilote peuvent être décrits en termes soit des schémas originaux élaborés pour le modèle simple, soit des schémas originaux auxquels on a apporté des compléments mineurs. Cette observation est vraie pour chacune des transcriptions de contrôle de terminal (élèves contrôleurs différents, même tâche). La conséquence la plus importante de cette constatation est que l'approche utilisée jusqu'ici pour modéliser le langage (i.e. l'utilisation de schémas) et la méthode employée pour élaborer les schémas semblent convenir. Cette constatation est d'autant plus significative que comme l'a mentionné au chapitre 6, la taille des échantillons est très réduite; cela témoigne donc de la stabilité des dialogues. Lors

de l'élaboration du tableau, on a aussi constaté que :

- Les schémas qui devront être ajoutés pour le contrôle de terminal seront à peu près identiques pour les deux transcriptions. Cela confirme que le dialogue est stable pour différents élèves contrôleurs.

- Le nombre de nouveaux schémas à ajouter est relativement restreint. Au premier abord, il semble que cinq nouveaux schémas devraient suffire à compléter un modèle suffisant pour les exercices de contrôle de terminal qui ont été observés.

On n'a pas apporté les modifications aux schémas originaux car ces derniers sont susceptibles d'évoluer, comme on l'a mentionné précédemment, lors d'une mise en oeuvre du modèle.

Par ailleurs, on a ébauché à l'annexe J le contenu des schémas additionnels qui sont nécessaires pour modéliser les exercices de terminal (pour l'élève contrôleur et le pseudo-pilote).

Enfin, on note que la définition des schémas additionnels peut se faire ici aussi de plusieurs façons : le découpage des schémas n'est pas unique. Par exemple, un des nouveaux schémas ébauchés (CFAD) pourrait être soit un nouveau schéma autonome, soit un nouveau schéma qui utilise certains des schémas du modèle préliminaire (ici, ALD et HDD). Il semble qu'une approche itérative utilisant l'ébauche des schémas additionnels pourrait être utilisée lors d'une simulation. Les résultats de la simulation ainsi que d'autres résultats sur l'optimisation des schémas (29) pourraient alors être utilisés pour raffiner le modèle.

8.2 Relâchement des contraintes sur la concaténation des messages et le nombre de destinataires d'une transmission

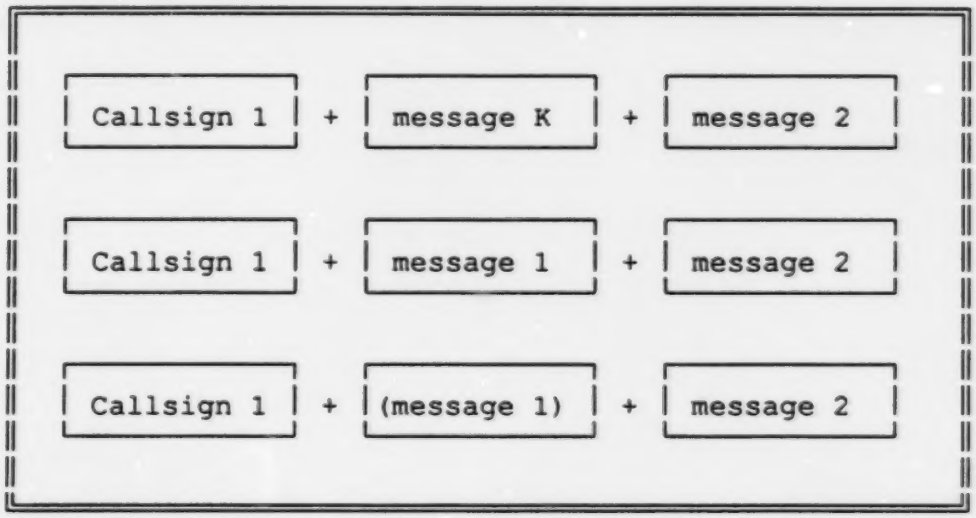
Jusqu'ici, on a négligé d'aborder une complexité qu'on rencontre dans les transcriptions : certaines transmissions soit véhiculent plus d'un message pour le même destinataire, soit véhiculent des messages différents à plus d'un destinataire. Cette question n'intéresse que les transmissions émanant des élèves contrôleurs, celles émanant des pseudo-pilotes ne posant pas de problèmes de modélisation. Les sections qui suivent

présentent le sommaire des résultats de l'analyse de ce genre de messages. L'analyse détaillée se trouve à l'annexe K.

8.2.1 Cas de plusieurs messages pour le même destinataire

Le tableau suivant présente le sommaire de l'analyse pour le cas où on a plusieurs messages pour le même destinataire.

Tableau XX Analyse des transmissions contenant plusieurs messages pour le même destinataire



On observe trois patrons principaux :

- Une transmission comporte l'indicatif d'appel du destinataire, un message de catégorie K et un second message.
- Une transmission comporte l'indicatif d'appel du destinataire, un message de catégorie quelconque et un second message.
- Une transmission comporte l'indicatif d'appel du destinataire, un message de catégorie quelconque et

un second message. Un processus de gestion d'erreur peut aussi se trouver dans la transmission (par exemple, une autocorrection d'un des messages de la transmission).

. Enfin, on a observé un maximum de deux messages différents par transmission. Il est toutefois possible qu'un nombre plus élevé de messages se retrouve dans une transmission si on obtient un échantillonnage plus large.

On modélise cette situation avec les règles suivantes :

. Règle : une transmission s'adresse à un seul destinataire et comporte plus d'une catégorie de message si a) il y a un seul indicatif d'appel à l'intérieur du message (sauf celui de l'émetteur du message) b) une seule catégorie de message n'explique pas de façon satisfaisante le contenu de la transmission.

. Règle : la réponse à un message émis par un appareil peut être une transmission débutant par un message correspondant au schéma "K" (le schéma K correspond à un accusé de réception) suivi d'un autre message pour le même appareil.

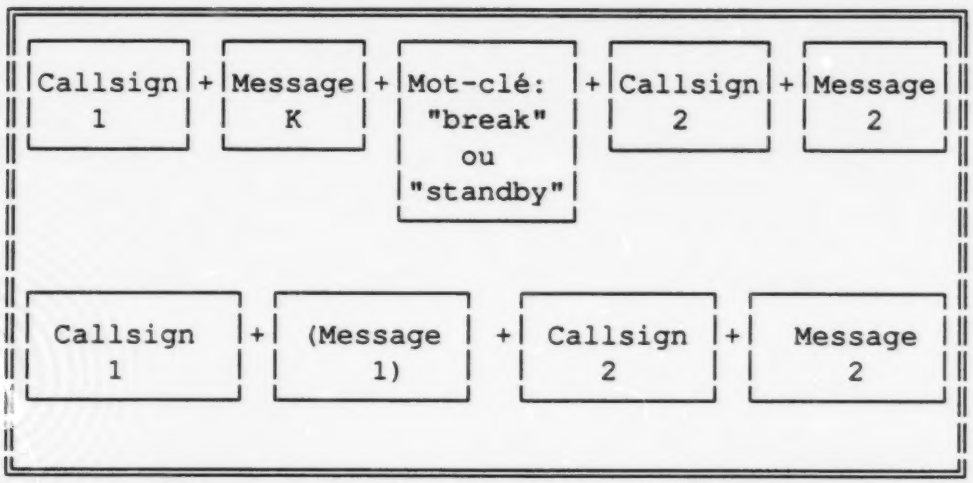
. Règle : une transmission comptant plus d'un message et envoyée à un seul destinataire peut aussi contenir un (ou plusieurs) processus de gestion d'erreur.

Ces règles sont assez grossières car elles ne tentent pas de définir le genre de messages qui sont susceptibles d'être concaténés dans une transmission : cela nécessiterait vraisemblablement une connaissance du détail de l'exercice ainsi que des connaissances approfondies en contrôle aérien.

8.2.2 Cas de messages différents pour plus d'un destinataire

Le tableau suivant illustre l'analyse du cas restant.

Tableau XXI Analyse des transmissions contenant des messages différents pour plus d'un destinataire



L'analyse des transcriptions montre qu'il y a un maximum de deux destinataires pour une transmission. Même si on pourrait soutenir qu'une même transmission pourrait intéresser plus de deux destinataires, cela semble très improbable. En effet, les possibilités d'erreurs de la part de l'élève contrôleur semblent trop élevées (il a à envoyer beaucoup d'information en peu de temps) et la gestion du dialogue devient très lourde (chaque directive entraîne un message de confirmation de chaque appareil). On considère donc qu'une transmission n'intéresse au plus que deux destinataires. On modélise cette situation avec :

. Règle : une transmission comporte plus d'un message et s'adresse à plus d'un destinataire si a) il y a répétition des éléments "Callsign" suivi de "message", pour des indicatifs d'appel différents, b) il y a la présence (facultative) d'un des mots-clés "break" ou "standby".

. Règle : il y a au plus deux destinataires pour une transmission.

Dès qu'on a reconnu la présence des éléments "Callsign" suivi de "message", on peut analyser chaque partie du message séparément.

8.3 Remarques sur la modélisation du langage

Jusqu'ici, on a montré que :

- On peut construire un modèle de langage basé sur des schémas pour un exercice simple de contrôle aérien.

- Les éléments du modèle simple (vocabulaire, schémas, etc.) peuvent servir à construire un modèle de langage pour un exercice complexe.

Comme on a pas considéré tous les types d'exercices ni toutes les situations de contrôle aérien qui sont utilisés à l'Institut, le modèle qui a été développé jusqu'ici n'est pas exhaustif. Par contre, il peut être utilisé comme base pour le développement d'un système de pseudo-pilote automatisé. Le modèle de langage pourra alors être amélioré au fur et à mesure qu'on acquérera de l'expérience avec le système. Il apparaît clair que l'amélioration du modèle nécessitera le concours d'une personne compétente en contrôle aérien. On conclut ainsi l'aspect de la modélisation du langage.

Le dernier objectif de cette étude est d'examiner l'utilisation des technologies de reconnaissance et de génération de parole en vue d'une utilisation dans un pseudo-pilote automatisé. C'est ce qu'on fait dans les chapitres qui suivent.

9. EXIGENCES ET POSSIBILITÉS POUR LA RECONNAISSANCE DE LA PAROLE

Cette section examine plusieurs aspects de la reconnaissance de la parole pour le pseudo-pilote automatisé. On définit d'abord sommairement le genre de reconnaisseur requis; on examine ensuite divers facteurs ergonomiques tels que l'apprentissage et l'effet de l'interaction avec une machine sur le dialogue.

9.1 Description des caractéristiques requises pour le reconnaisseur de parole

Si on désire obtenir un pseudo-pilote automatisé qui ait un comportement vocal semblable à celui d'un humain, i.e. qui soit transparent pour l'élève contrôleur, on doit avoir les caractéristiques suivantes :

- Reconnaissance de la parole connectée, comme dans le cas du pseudo-pilote actuel.
- Temps de réaction du pseudo-pilote (incluant le temps requis par le reconnaisseur de parole) comparable à celui d'un pseudo-pilote humain. On a évalué grossièrement ce temps de réaction à 1.75 secondes[1].
- Activation par un interrupteur commandé au pied, comme pour le poste de travail actuel.

---

[1] Cette valeur correspond au délai moyen entre l'émission d'un message par l'élève-contrôleur et la réponse par le pseudo-pilote. On a obtenu cette valeur en effectuant quatre mesures, au hasard, pour chacun des quatre enregistrements utilisés dans cette étude. Les valeurs obtenues (arrondies à la seconde) pour chaque enregistrement sont respectivement 3, 2, 1, 3, 1, 1, 1, 1, 1, 1, 2, 3, 1, 1, 2, et 4. Le délai moyen est donc 1.75 secondes.

- Vocabulaire d'environ 200 mots : ce nombre correspond au vocabulaire total pour tous les exercices examinés ici. On peut toutefois recourir à un vocabulaire plus petit pour certains exercices si on prévoit un vocabulaire adapté à chaque exercice.

- Selon la stratégie adoptée au niveau de la syntaxe (stricte, évolutive ou libre), il faut prévoir des capacités différentes en termes de nombre de branchements permis par l'automate à états finis du reconnaisseur de parole. Dans le cas d'une syntaxe libre, ce problème ne se pose pas. Par ailleurs, pour une syntaxe évolutive, il faut s'assurer que le reconnaisseur permet de commuter rapidement entre diverses syntaxes.

- Le reconnaisseur de parole doit aussi pouvoir analyser une phrase comptant au minimum de 32 mots (pour un exercice simple) à 48 mots (pour un exercice complexe).

Le système doit aussi être relativement insensible aux conversations qui ont lieu dans les postes adjacents, ce qui suggère l'emploi d'un microphone directionnel à annulation de bruit. Compte tenu du comportement de certains reconnaisseurs de parole, le même microphone devrait être utilisé pour l'apprentissage d'un reconnaisseur de parole et pour l'exploitation du reconnaisseur. De plus, les microphones directionnels sont par définition sensibles à leur orientation ce qui peut entraîner des variations de niveau de signal appréciables (30). On ignore encore si ce problème se présentera dans le pseudo-pilote automatisé, mais le cas échéant, il peut exister des solutions :

- Action au niveau du port du casque d'écoute : on peut vraisemblablement encourager un port "correct" et standardisé du casque d'écoute chez les élèves contrôleurs lors de l'entraînement sur simulateur.

- Action au niveau de l'élocution : on peut d'abord encourager une élocution claire et distincte de la part de l'élève contrôleur, par exemple lors d'une familiarisation avec le reconnaisseur de parole. Lors de l'exploitation régulière, on peut recourir à une méthode additionnelle : on peut afficher sur un indicateur le niveau sonore perçu par le reconnaisseur

de parole et laisser à l'élève contrôleur le soin de contrôler son élocution en conséquence. Cette dernière méthode a l'inconvénient de mobiliser une partie de l'attention de l'élève contrôleur à seule fin d'aider le reconnaisseur de parole. Cette solution semble peu acceptable (la tâche de l'élève contrôleur est d'apprendre le contrôle aérien, pas de se faire entendre de la machine).

- Action sur le gain d'entrée du reconnaisseur : certains reconnaisseurs de parole incorporent un contrôle automatique du gain du signal d'entrée (ITT VRS 1280, SSI Phonetic Engine, Marconi ASR 1000, etc.). Typiquement, le gain est ajusté après qu'une phrase complète ait été prononcée. Cette méthode (combinée à un port du casque correct) semble être celle qui perturbera le moins l'élève contrôleur.

Il serait idéal d'utiliser un reconnaisseur de parole qui serait multilocuteur. Toutefois, nous ne possédons pas assez d'information sur des reconnaisseurs de parole multilocuteurs utilisant de la parole connectée, pour la taille de vocabulaire considéré, pour conclure. Nous proposons donc l'utilisation d'un reconnaisseur de parole monolocuteur, ce qui nous semble susceptible de fournir le meilleur taux de reconnaissance. Cela nécessite donc de faire l'apprentissage du reconnaisseur de parole, i.e. de fournir des échantillons de voix de chaque usager au reconnaisseur de parole. Il est donc important que l'appareil choisi dispose d'outils de création des échantillons de parole qui soient conviviaux. Pour améliorer la qualité des échantillons, il importe aussi que l'apprentissage multi-passes soit possible.

9.2 Apprentissage avec le reconnaisseur de parole

Un des facteurs qui semble prépondérant dans l'application réussie de la reconnaissance de la parole est l'attitude de l'usager (31) (32). Le premier moment où la plupart des usagers prennent contact avec un reconnaisseur de parole est lors de la phase d'apprentissage. Il est donc très important de tirer partie de la phase d'apprentissage pour susciter et encourager la coopération de l'usager, en plus évidemment d'obtenir des échantillons de voix qui soient de bonne qualité pour le reconnaisseur de parole. On propose ici un cadre d'apprentissage

qui pourrait permettre d'atteindre ces objectifs :

- D'abord une phase où l'élève contrôleur se familiarise avec le reconnaisseur de parole, avec un vocabulaire réduit. Les meilleurs résultats en reconnaissance sont généralement obtenus pour des usagers qui ont eu le temps de se familiariser avec le reconnaisseur de parole (31). De plus, diverses expériences suggèrent d'y aller progressivement au point de vue du vocabulaire. Cette phase initiale d'apprentissage pourrait être, par exemple, à l'occasion d'une session où l'élève contrôleur se familiarise avec certaines parties de la phraséologie utilisée en contrôle aérien comme le code d'épellation en radio-téléphonie de l'OACI (alpha..zulu), les nombres, etc. On est en mode de mots isolés.

- Ensuite une phase intermédiaire où on complète le vocabulaire et où on introduit la parole connectée.

- Enfin, une phase où on fait l'apprentissage de la phraséologie en mots connectés.

On peut facilement imaginer d'autres scénarios; le point important est de considérer l'apprentissage comme une opportunité à exploiter et non comme un obstacle à surmonter. Une possibilité intéressante serait d'intégrer l'apprentissage du reconnaisseur de parole à l'apprentissage de certaines parties de la phraséologie. Cela pourrait être utile pour l'apprentissage de la phraséologie en général et pour l'apprentissage du reconnaisseur de parole en particulier. Ce point a d'ailleurs été soulevé lors de discussions avec des constructeurs de systèmes (9) : ces derniers pensent qu'il pourrait être profitable de familiariser les élèves contrôleurs avec la technologie vocale par le biais d'un système simple (par exemple un "jeu" où l'élève guiderait par la voix un témoin lumineux le long d'un parcours donné).

Il y a d'autres facettes qui sont relatives à l'apprentissage et qui peuvent contribuer à une bonne performance :

- On obtient les meilleurs résultats quand l'apprentissage est supervisé et qu'il est réalisé dans un environnement acoustique le plus près possible

de l'environnement dans lequel l'exploitation aura lieu (32) (33).

. L'apprentissage devrait être fait de façon à reproduire le ton qui sera utilisé par l'élève contrôleur lors de l'exploitation du reconnaisseur de parole; un usager articule différemment selon qu'il lit un texte ou qu'il commande un système (34).

9.3 Effet de l'interaction avec une machine sur le comportement langagier

On sait qu'un usager ne converse pas de la même façon selon qu'il pense s'adresser à un humain ou à une machine. Dans le cadre d'études sur le dialogue entre des humains et des machines (35), on a observé que l'interaction avec une machine peut induire les comportements suivants chez l'humain :

. Le dialogue est stabilisé, i.e. les communications ont tendance à se rapprocher d'une forme "standard".

. Le dialogue a tendance à être moins interactif.

. L'usager a tendance à imiter le ton et le rythme lent et monocorde de la machine.

Il est difficile de déterminer précisément quel facteur (l'appréhension de ne pas être compris, l'effet d'une voix synthétique, etc.) tend à susciter ce comportement; il reste que ce comportement est favorable à la reconnaissance de la parole. Même si cette étude a été effectuée dans un autre contexte (la demande d'horaire de trains), il est raisonnable de croire qu'on pourrait faire les mêmes observations dans le cas d'un pseudo-pilote automatisé. Il pourrait donc être profitable de chercher à susciter ce comportement langagier "stabilisé", par exemple en expliquant convenablement aux usagers les limites de compréhension d'un pseudo-pilote automatisé éventuel.

Sans prétendre être exhaustif, ce chapitre a souligné plusieurs aspects de la reconnaissance de la parole qui doivent être pris en compte pour faciliter l'exploitation d'un pseudo-pilote automatisé. Le prochain chapitre examine la génération de la parole et propose un choix qui tient compte de la tâche, des technologies disponible et des facteurs ergonomiques.

10. EXIGENCES ET POSSIBILITÉS POUR LA GÉNÉRATION DE LA PAROLE

Cette section vise à identifier une méthode de génération de parole qui soit appropriée pour un système de pseudo-pilote automatisé. On examine d'abord les particularités pertinentes de la tâche. On décrit ensuite sommairement deux méthodes de génération de parole parmi les plus usuelles. On examine ensuite les aspects ergonomiques de la voix synthétique selon deux points de vue : d'abord sa perception, ensuite l'effet de son utilisation sur le comportement langagier d'un usager. Enfin, on choisit une méthode en tenant compte de la tâche, des méthodes de génération de parole et des aspects ergonomiques.

10.1 Particularités pertinentes de la tâche

La première constatation que l'on fait en examinant les transcriptions est que dans le simulateur utilisé actuellement au centre de formation de Transports Canada à Cornwall, l'élève contrôleur n'a affaire en général qu'à un même individu lors d'un exercice donné. Ce dernier joue le rôle des différents pilotes en plus de jouer le rôle de contrôleur d'autres zones de contrôle. On n'utilise donc qu'une seule voix par élève contrôleur et par exercice dans le système actuel.

Comme on l'a montré précédemment (chapitre 6), les messages émis par les pseudo-pilotes sont assez stéréotypés, bien qu'à un degré moindre que pour les élèves contrôleurs. On y observe un certain nombre de variations : ces variations tiennent surtout aux abréviations qui varient entre les différents pseudo-pilotes (les pseudo-pilotes peuvent changer entre les exercices) et à l'emploi de formules de politesse (par exemple, "good day" ou "good afternoon"). On a déjà établi des schémas (voir Annexe D) qui correspondent aux messages émis par les pseudo-pilotes, lors de la mise au point du modèle préliminaire du dialogue.

On établit ensuite la taille et la composition du vocabulaire requis pour la tâche de pseudo-pilote en utilisant la même méthode que dans le cas du vocabulaire des élèves contrôleurs. Le tableau qui suit montre les résultats de cette

analyse pour toutes les transcriptions (i.e. exercices simples et complexes) qui ont été utilisées jusqu'ici. La liste complète des mots se trouve à l'annexe L.

Tableau XXII Vocabulaire des pseudo-pilotes pour les quatre transcriptions

Exercice simple (basse altitude)
(Transcriptions 1a et 2a combinées)

Taille du vocabulaire d'usage général	Nombre d'éléments standards	Taille du vocabulaire total
56 mots	79 mots	35 mots

Exercice complexe (terminal)
(Transcriptions 3a et 4a combinées)

Taille du vocabulaire d'usage général	Nombre d'éléments standards	Taille du vocabulaire total
86 mots	79 mots	65 mots

En prenant une marge de sécurité de 25 % (i.e. comme pour le vocabulaire des élèves contrôleurs), les vocabulaires requis par les pseudo-pilotes lors des exercices simples et complexes sont de 169 mots et 206 mots (i.e. 135 * 1.25 et 165 * 1.25), respectivement. De plus, comme dans le cas des élèves contrôleurs, le détail des analyses montre que les pseudo-pilotes utilisent à peu de choses près le même vocabulaire, même si la variabilité dans le vocabulaire utilisé entre les pseudo-pilotes est plus grande qu'entre les élèves contrôleurs.

10.2 Description et caractéristiques des méthodes de génération

On décrit ici très sommairement les deux méthodes les plus usuelles pour générer des paroles de façon automatique : la

concaténation de mots ou segments de phrases et la synthèse à partir du texte. On veut identifier les caractéristiques fonctionnelles pertinentes de ces méthodes par rapport à une application dans un pseudo-pilote automatisé.

10.2.1 Description et caractéristiques de la concaténation de mots ou de segments de phrases

Il s'agit d'enregistrer les éléments (mots ou segments de phrases) des messages à générer et de les restituer sur demande en les concaténant dans un message (36). On peut ainsi créer un message de toutes pièces, mot par mot, ou encore utiliser une phrase fixe à laquelle on ajoute des éléments variables (style "horloge parlante"). La figure qui suit montre les étapes principales qui permettent d'obtenir les mots ou segments de phrases qui seront utilisés pour la concaténation :

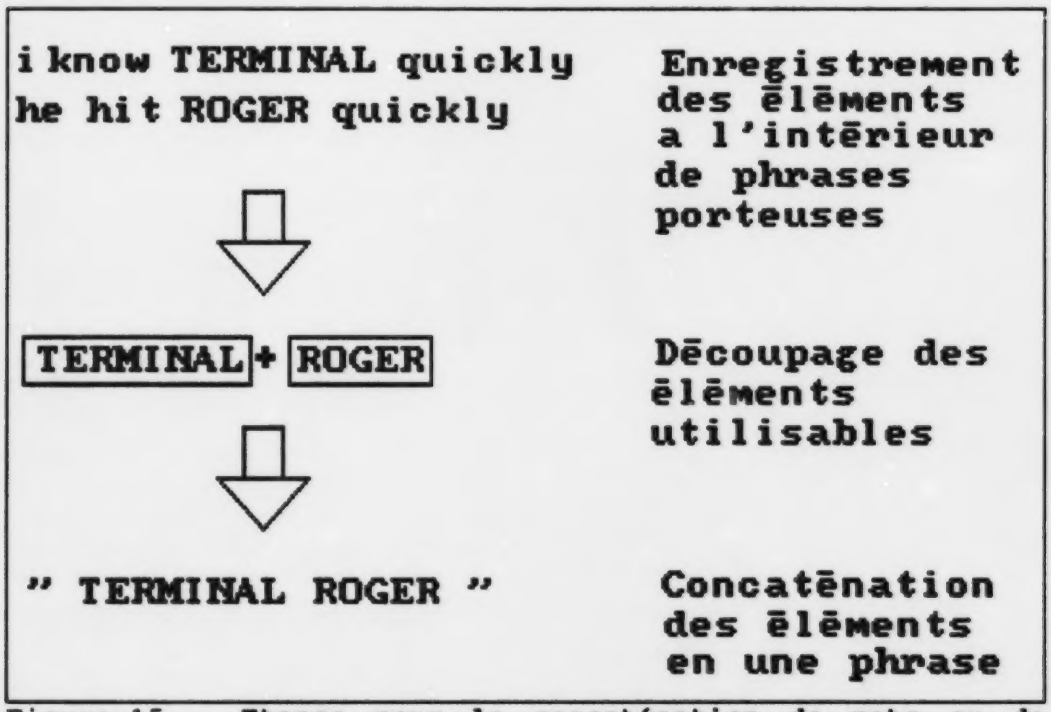


Figure 15 **Etapes pour la concaténation de mots ou de segments de phrases**

Pour utiliser cette méthode, on enregistre d'abord chacun des mots ou des segments de phrase (i.e. un ensemble cohérent de mots) à l'intérieur d'une "phrase porteuse" : on veut ainsi obtenir un échantillon qui soit neutre (i.e. sans intonation particulière) et sans effet de coarticulation avec d'autres éléments. L'échantillon obtenu est ensuite numérisé et stocké soit sous forme brute (i.e. sans compression), soit sous forme codée (i.e. avec compression). La compression vise à réduire l'espace requis pour stocker les échantillons : il existe des algorithmes de codage qui sont standardisés et qui permettent d'obtenir une réduction importante des besoins en stockage sans entraîner de dégradation appréciable de la qualité de la parole (37). On apprécie mieux l'importance de la compression avec un exemple. Si on accepte une bande passante de 5 KHz pour la voix qui sera utilisée et que l'on numérise cette voix avec un convertisseur analogique à numérique de 12 bits par échantillon, on obtient :

5 KHz * 2 échant./cycle * 12 bits/échant. = 120 000 bits/sec

Cela correspond à 15 K-octets (120 000 bits / (8 bits / octet)) de stockage sur un ordinateur pour chaque seconde de parole enregistrée. A ce rythme, on consomme un peu moins de 1 méga-octet par minute de parole enregistrée. En comparaison, un codage du signal numérisé utilisant un algorithme ADPCM-32 Kbits/sec ramène les besoins en stockage à 4 kilo-octets par seconde de parole enregistrée, d'où une économie de stockage de près de 4 à 1.

Dans le cas de la concaténation de mots ou de segments de phrases, l'intelligibilité de la voix résultante est généralement bonne (38). Par contre, comme les mots ou segments de phrases sont neutres, la voix concaténée est prosodiquement assez différente de la voix naturelle. Des recherches ont donc été faites pour tenter d'améliorer cet aspect de la méthode (39) (40) : il existe aujourd'hui des façons de générer de la voix concaténée en tenant compte de certains éléments de prosodie. On manque toutefois d'évaluations sur l'impact de ces méthodes sur la qualité de la voix. De plus, ces modifications entraînent des manipulations supplémentaires pour générer la voix concaténée.

Il reste à considérer l'aspect de la souplesse ou flexibilité de la concaténation de mots ou de segments de

phrases. En effet, à la différence des autres méthodes où un mot est bâti à partir d'unités plus petites (phonèmes, diphones, etc), on ne peut concaténer ici que les mots ou segments de phrase qui ont déjà été enregistrés. Une extension du vocabulaire requiert l'ajout de mots ou segments de phrase aux unités déjà stockées, en utilisant de préférence le même locuteur (il serait bizarre de retrouver des voix différentes à l'intérieur de la même phrase). Comme chaque mot ou segment de phrase doit être stocké comme une entité distincte, on conçoit que cette méthode ne convienne que pour des vocabulaires stables (i.e. qui ne soient pas susceptibles de changer beaucoup) et peu étendus.

10.2.2 Description de la synthèse à partir du texte

Comme son nom l'indique, la synthèse à partir du texte permet d'entrer des mots et des phrases dans un système et d'en obtenir une voix synthétique représentant ces mots et phrases. La figure suivante illustre les composantes principales d'un système complet de synthèse à partir du texte (adapté de (41)).

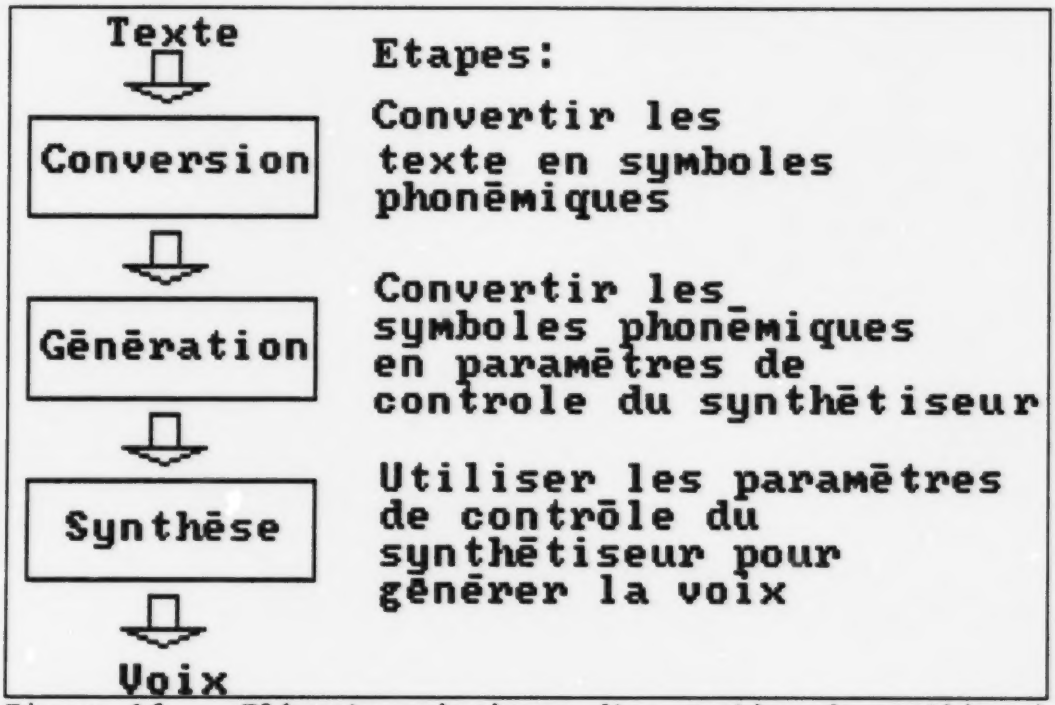


Figure 16 **Eléments principaux d'un système de synthèse à partir du texte**

Comme ces systèmes sont appelés à accepter des mots ou phrases quelconques, il n'est pas viable d'enregistrer tous les mots que l'on est susceptible de rencontrer. On commence par "traduire" les mots en éléments acoustiques "fondamentaux" plus petits que les mots, i.e. des phonèmes, pour générer une voix synthétique. Un phonème est le plus petit son qui ait une identité propre dans une langue. Il y a un nombre restreint de phonèmes dans chaque langue; par exemple, on en compte environ 40 en anglais.

Pour générer de la parole synthétique, on doit donc manipuler ou assembler les phonèmes de façon satisfaisante. L'idée de concaténer les phonèmes composant chaque mot est alors séduisante. Il y a toutefois un obstacle majeur : ce ne sont pas les phonèmes seuls qui portent l'information dans la parole, mais plutôt les transitions entre les phonèmes. En effet, la prononciation d'un son dépend, entre autres, des sons précédents (qu'on a prononcés) et des sons à venir (qu'on se prépare à prononcer). La prononciation est aussi affectée par

l'intonation et le débit.

Plusieurs approches permettent d'assembler les phonèmes en tenant compte des transitions entre eux. La plus populaire consiste à modéliser ces transitions en utilisant de règles. Cette méthode a comme caractéristiques de :

- Utiliser une quantité minime de mémoire pour stocker toutes les unités constituantes (phonèmes) des sons d'une langue.

- Nécessiter un grand nombre de règles pour représenter les transitions entre les phonèmes. La qualité de la synthèse dépend de la forme et du nombre de règles du système.

Comme les phonèmes varient d'une langue à une autre, on réalise qu'un système de synthèse par règles fonctionne mieux dans la langue pour laquelle il a été développé.

10.3 Aspects ergonomiques : perception de la parole générée

On examine ici divers aspects ergonomiques qui peuvent guider le choix entre les méthodes de génération de parole qui ont été présentées. On s'intéresse surtout à :

- La "performance" de diverses voix synthétiques : cet aspect englobe des indicateurs traditionnels sur l'évaluation de la voix synthétique tels que la netteté, l'intelligibilité, le naturel, l'acceptabilité, etc.

- La charge cognitive du traitement d'une voix synthétique, comparativement à une voix naturelle.

Ces deux aspects sont surtout intéressants à cause du lien direct qui les rattache à la tâche de l'élève contrôleur. En effet, l'intelligibilité de certaines informations critiques (par exemple l'altitude, le cap, etc.) est de première importance. De la même façon, on veut aussi utiliser une voix qui soit près de la voix d'un pseudo-pilote humain, donc de bonne qualité subjective. Pour la charge cognitive, il est justifié de ne pas vouloir rendre la tâche de l'élève contrôleur plus difficile qu'elle ne l'est déjà en utilisant une voix

synthétique qui soit peu appropriée.

10.3.1 Performance de diverses voix synthétiques

On examine d'abord certains facteurs objectifs d'évaluation de la qualité de la voix synthétique en faisant d'abord les distinctions suivantes (42) :

- Netteté : concerne les mesures avec des symboles dénués de sens (note : dans plusieurs études en langue anglaise, on réfère à la netteté comme étant l'intelligibilité segmentale (segmental intelligibility)).

- Intelligibilité : se détermine avec des mots ou des phrases ayant un sens, supports ou véhicules d'une idée qu'il faut comprendre.

Sans entrer dans le détail des méthodologies de test de la netteté, on constate qu'il est facile de retrouver plusieurs études qui rapportent les résultats d'évaluation pour des systèmes de synthèse à partir du texte. Le tableau suivant illustre certains de ces résultats (adapté de (43)) :

Tableau XXIII Sommaire de résultats sur des tests de netteté, pour des systèmes de synthèse à partir du texte.

Système de synthèse partir du texte	Taux d'erreur global, test MRT, en %
Voix naturelle	0,53
Dectalk (Paul)	3,25
Dectalk (Betty)	5,72
Prose 3.0	5,72
MITalk-79	7,00
Infovox SA 101	12,50
Berkeley	14,14
TSI-proto 1	17,75
Votrax Type'n'Talk	27,44
Echo	35,56

On remarque d'abord que toutes les voix synthétiques évaluées sont pires que la voix témoin (voix naturelle). On voit aussi que c'est une des voix du système Dectalk (Paul) qui offre les meilleures performances.

Ces résultats sur la netteté et l'intelligibilité sont incomplets : on y néglige les effets de l'entraînement aux différentes voix synthétiques. De fait, l'accoutumance à une voix synthétique peut avoir un effet marqué sur la netteté et l'intelligibilité perçues de cette voix (44).

On n'a pas pu retracer de résultats équivalents dans le cas de systèmes utilisant la concaténation de mots ou de segments de phrases. Il ne semble pas y avoir non plus d'études

comparatives entre les systèmes de synthèse à partir du texte et les systèmes utilisant la concaténation de mots ou de segments de phrases.

On a mentionné la différence entre la netteté et l'intelligibilité. Cette dernière tient compte du contexte, i.e. des éléments linguistiques et pragmatiques impliqués dans le message (45). La familiarité avec la tâche et, dans le cadre de cette étude, avec une phraséologie donnée peut affecter considérablement l'intelligibilité d'un message donné. On a montré que, pour une voix synthétique donnée, des usagers familiers avec la phraséologie utilisée en contrôle aérien obtenaient de meilleurs résultats que des usagers familiers avec un autre type de phraséologie (46).

Même si la mesure de la netteté et de l'intelligibilité de la voix synthétique sont des éléments fondamentaux, on doit leur adjoindre une dimension subjective qui tienne compte de l'auditeur humain. On distingue ainsi :

- Le naturel, i.e. le degré de ressemblance d'une voix artificielle avec une voix humaine.

- L'acceptabilité, i.e. la tolérance d'un auditeur humain pour une voix artificielle.

Comme il existe plusieurs études qui évaluent un ou plusieurs des critères qui viennent d'être définis, on peut se demander quel(s) critère(s) on doit privilégier pour évaluer et choisir une voix synthétique. Cette tâche est rendue plus facile quand on reconnaît les liens qui existent entre certains des critères. En effet, on a montré qu'il y a une forte corrélation entre l'acceptabilité d'une voix synthétique et sa netteté perçue (47), (48). Même si ces résultats sont incomplets (on ignore comment se comporte l'acceptabilité pour deux voix de même netteté), on peut croire qu'on a avantage à choisir, parmi des voix différentes, la voix ayant la meilleure netteté. Cela est d'autant plus vrai que dans un pseudo-pilote automatisé, on peut vouloir simuler pour des fins de formation un canal de communication radio. Dans ce cas, il est possible que les différences en intelligibilité soient amplifiée.

En conclusion de cette section, on constate que :

- Pour les systèmes de synthèse à partir du texte, le système Dectalk avec la voix "Paul" obtient en général

les meilleurs résultats.

- Il ne semble pas y avoir d'études comparatives de performance entre les systèmes de synthèse à partir du texte et les systèmes utilisant la concaténation de mots ou de segments de phrases.

- Il semble exister une corrélation positive entre la netteté et l'acceptabilité d'une voix.

Au point de vue perceptuel, il reste à voir l'impact de l'utilisation d'une voix synthétique sur la charge cognitive d'un usager. C'est ce que l'on fait à la prochaine section.

10.3.2 Charge cognitive du traitement d'une voix synthétique

En général, on s'entend pour dire que la performance des voix synthétiques en terme de netteté et d'intelligibilité est inférieure à celle d'une voix humaine (49). Mais au-delà d'entendre quelque chose, il faut comprendre ce quelque chose : on fait donc une distinction entre la perception et la compréhension. De plus, l'effort de compréhension ne doit pas nuire à l'accomplissement de la tâche principale dans laquelle est engagé l'usager.

Il existe relativement peu de données sur la compréhension et l'utilisation de l'information acquise par l'intermédiaire d'une voix synthétique. Certains résultats semblent toutefois indiquer que :

- Il est plus difficile de se rappeler et d'utiliser des informations reçues par le biais d'une voix synthétique (50) (51). On croit que la perception de parole synthétique requiert une capacité de traitement accrue de la part des ressources cognitives d'un usager pour encoder cette voix.

- Le temps requis pour analyser et utiliser l'information provenant d'une voix synthétique est plus important que dans le cas d'une voix naturelle, et cela d'autant plus que la voix synthétique est de mauvaise qualité (43).

Il faut noter que ces études n'ont pas cherché à quantifier de

façon stricte la qualité des voix de synthèse utilisées. Il existe aussi peu de comparaisons sur la compréhension de plusieurs voix synthétiques (ou plutôt artificielles) entre elles. Dans un cas cependant (52), on a comparé une voix naturelle, des voix numérisées (à 9.6 kbs et à 2.4 kbs) et une voix de synthèse (Dectalk). On a pas relevé de différence en résultats de compréhension entre la voix naturelle et la voix numérisée de meilleure qualité (9.6 kbs); on a par contre trouvé que les résultats de compréhension étaient moins bons dans le cas de la voix numérisée de moins bonne qualité (2.4 kbs) et dans le cas de la voix de synthèse (Dectalk).

Comme on l'a mentionné quand on a traité d'intelligibilité, il est possible que l'habitude et l'entraînement améliorent la perception (l'intelligibilité) de la voix synthétique. Les opinions divergent toutefois sur le fait de savoir si cette amélioration de la perception se traduit par une réduction des besoins en capacité de traitement de la part de l'opérateur. A partir de éléments de cette sections, on conclut que :

. Les résultats de tests de compréhension et de rétention sont meilleurs pour les voix qui présentent la meilleure intelligibilité.

10.4 Effets de l'utilisation de parole générée artificiellement sur le comportement langagier

On examine ici les impacts de deux effets de l'utilisation de parole générée artificiellement sur le comportement langagier d'un usager :

. Un effet de bas niveau : l'adaptation.

. Un effet de haut niveau : la stabilisation du dialogue.

Les termes bas et haut niveau décrivent approximativement la complexité des mécanismes cognitifs impliqués.

10.4.1 L'adaptation

L'adaptation est un facteur bien connu en psychologie

sociale. En effet, on sait qu'il est fréquent qu'un interlocuteur adapte spontanément son langage à celui de son interlocuteur : on parle de convergence quand un individu rend son langage semblable à celui de son interlocuteur, et de divergence quand il l'en différencie (22). En général, on associe la convergence avec une attitude polie et coopérative alors que la divergence accentue les différences sociales ou de groupe.

L'adaptation peut agir à plusieurs niveaux, à partir de la construction des phrases jusqu'au choix et au mode de prononciation des mots.

On a aussi noté que certains locuteurs tendent à imiter l'invite vocale fourni par une voix synthétique lorsqu'ils procèdent à l'apprentissage de leur voix (acquisition d'échantillons de leur voix) sur un reconnaisseur de parole (45).

10.4.2 La stabilisation du dialogue

On a mentionné dans un chapitre précédent que l'interaction avec une machine peut conduire à une stabilisation du dialogue. Par ailleurs, une des tâches qu'accomplit l'interface (ici le dialogue entre l'élève contrôleur et le pseudo-pilote automatisé) est de susciter chez l'usager une représentation du système auquel il a affaire (53). On discerne alors un des rôles que pourra avoir à jouer la voix synthétique, i.e. :

- Susciter chez l'usager la perception qu'il dialogue avec un système capable de comprendre et d'interpréter ses directives, mais doté d'une capacité de compréhension du langage restreinte à une forme d'expression assez stricte à la fois pour la structure des phrases, les mots utilisés et la prononciation.

On propose donc qu'il peut être possible d'exploiter certains aspects de la voix synthétique pour rappeler constamment à l'usager qu'il doit "surveiller son langage" car il n'a pas affaire à un humain. Certains chercheurs commencent d'ailleurs à s'intéresser à cet aspect des interfaces utilisant la reconnaissance et la génération de la parole (54). Par exemple, on pourrait générer les réponses du pseudo-pilote avec une voix de haute qualité mais en accentuant la conformité à la

phraséologie officielle, au-delà de ce que le pseudo-pilote humain fait actuellement. En effet, on a noté qu'il est fréquent que le pseudo-pilote humain introduise une certaine variabilité dans ses messages. Le fait de générer des messages respectant à la lettre la phraséologie standard pourrait aider à la stabilisation du dialogue en rappelant constamment à l'élève contrôleur qu'il a affaire à une machine. Il y aurait peut-être aussi un gain pour l'apprentissage de la phraséologie, de la part de l'élève contrôleur.

10.5 Choix d'une voix synthétique pour le pseudo-pilote automatisé

Cette section ne vise pas à proposer un choix définitif et précis d'une voix artificielle pour le pseudo-pilote automatisé. On propose plutôt une démarche qui tient compte des éléments qui ont été identifiés jusqu'ici et qui pourra être adaptée au fur et à mesure du développement d'un prototype de pseudo-pilote automatisé.

10.5.1 Exigences et possibilités des méthodes disponibles

On désire une voix artificielle qui :

- Permette des modifications éventuelles faciles du vocabulaire ou des réponses du pseudo-pilote automatisé.
- Possède une grande intelligibilité pour favoriser un dialogue exempt d'erreurs, éviter d'augmenter la charge de travail de l'élève contrôleur et susciter la plus grande acceptabilité possible du système.
- Suscite un comportement langagier favorable chez l'élève contrôleur en lui laissant percevoir constamment qu'il a affaire à un système compétent dans un domaine restreint du langage naturel.
- Permettre de générer des voix différentes pour représenter les différents pseudo-pilotes auxquels un élève contrôleur doit avoir affaire dans un exercice.

. Puisse s'adapter à un vocabulaire stable et à un ensemble fermé de réponses du pseudo-pilote automatisé.

L'approche la plus flexible et la plus facile au point de vue des modifications éventuelles au vocabulaire et aux réponses du pseudo-pilote est la synthèse à partir du texte. Cette approche a aussi le mérite de laisser clairement savoir à l'élève contrôleur qu'il a affaire à une machine. Malheureusement, le meilleur système de synthèse à partir du texte disponible (Dectalk avec la voix de Paul) ne possède pas une intelligibilité aussi bonne que la parole naturelle : il y a donc des conséquences possibles au niveau de la compréhension, de la charge de travail de l'élève contrôleur et de l'acceptabilité. Cela est d'autant plus vrai si l'on utilise les autres voix disponibles sur le système. Un dernier inconvénient est que le Dectalk est restreint à une utilisation en anglais et ne pourrait être utilisé pour un pseudo-pilote opérant en français (le cas échéant, on pourrait toutefois tenter d'utiliser un synthétiseur à partir du texte conçu pour fonctionner en français).

L'autre approche est celle de la concaténation de mots ou de segments de phrases. Les modifications éventuelles de vocabulaire ou de réponses du pseudo-pilote sont moins faciles : chaque nouveau mot ou segment de phrases doit être enregistré. Par ailleurs, il est possible de produire autant de type de voix que désiré (ce qui demande toutefois une quantité additionnelle de capacité de stockage). Ce point est important pour ajouter au réalisme de la formation de l'élève contrôleur : on pourrait ainsi donner un accent approprié aux voix associées à chaque indicatif d'appel (par exemple, accent italien pour Lufthansa, accent britannique pour British Airways, etc.). De plus, avec un découpage approprié des phrases, on peut probablement obtenir une intelligibilité globale (i.e. au niveau de chacune des phrases) proche de celle de la voix naturelle. La voix obtenue est cependant déficiente au point de vue de la prosodie. Il y a actuellement des recherches en cours pour améliorer le contour prosodique de la voix concaténée; il est déjà possible d'obtenir des phrases concaténées qui sont difficilement discernables de la voix naturelle. Cette amélioration nécessite toutefois du travail supplémentaire pour générer les phrases.

Notons enfin que cette approche est celle qui a été retenue par les firmes qui s'occupent de la partie vocale des pseudo-pilotes qui sont développés actuellement (ITT, Verbex, Speech

Systems Inc.).

10.5.2 Démarche proposée

On propose la démarche suivante :

- Générer un vocabulaire et un ensemble de réponses à partir de concaténation de mots et de segments de phrases pour un certain nombre de voix, sans correction pour la prosodie.

- Mettre à l'épreuve le vocabulaire et l'ensemble de réponses sur un prototype. Evaluer l'acceptabilité, l'intelligibilité et l'effet sur la charge de travail par rapport à une voix naturelle (lors d'une simulation avec compère), pour des exercices de simulation.

A partir de ces résultats, il sera possible de conclure soit à la bonne performance de la voix concaténée "simple", soit à la nécessité d'évaluer la faisabilité d'une voix concaténée avec prosodie. Dans ce dernier cas, il faudra évaluer les améliorations espérées avec les désavantages possibles (traitement additionnel requis).

11. CONCLUSION

Cette étude a d'abord permis d'identifier les principales organisations qui étaient impliquées dans le développement de pseudo-pilotes automatisés, tant en Amérique du Nord qu'en Europe.

On a ensuite élaboré un modèle de langage opératif qui est adapté à divers types d'exercices de simulation utilisés pour la formation des élèves contrôleurs au Canada. On a ainsi proposé une approche qui permet de développer un système adapté à des exercices simples, et qui possède aussi tout le potentiel voulu pour être utilisé pour des exercices plus complexes.

Par la suite, on a présenté diverses recommandations sur l'emploi de technologies de reconnaissance et de génération de la parole qui tiennent compte à la fois de la tâche à accomplir, des capacités de la technologie et de l'ergonomie du dialogue.

Enfin, l'examen des travaux menés ailleurs et des exercices qui ont pu être observés à l'Institut de formation de Transports Canada à Cornwall, ainsi que le travail de modélisation du langage opératif nous permettent de croire que :

- La réalisation d'un pseudo-pilote automatisé qui permet de fournir une interaction réaliste à l'élève contrôleur peut se faire au moins dans le cadre d'exercices simples, pour des élèves contrôleurs débutants.

- Pour obtenir une interaction entre l'élève contrôleur et le pseudo-pilote automatisé qui soit similaire à celle que l'on retrouve aujourd'hui, on doit probablement recourir à un modèle de langage basé sur des schémas. Ce modèle doit aussi probablement contenir des connaissances sur la tâche, sur l'évolution du dialogue, etc.

- On a vraisemblablement avantage à incorporer l'obtention des échantillons de voix requis par le reconnaisseur de parole dans l'apprentissage de la phraséologie proprement dite. Par exemple, il peut

être possible d'utiliser le reconnaisseur de parole pour enseigner certains aspects élémentaires de la phraséologie.

Toutefois, au delà de la faisabilité technique d'un tel projet, nous croyons qu'il y a d'autres questions qui doivent être résolues, et ce préalablement au développement d'un pseudo-pilote automatisé.

D'abord, il semble primordial de déterminer l'impact souhaité d'un pseudo-pilote sur la formation des élèves contrôleurs. Souhaite-t-on seulement leur fournir un système qui répondra à leurs commandes, ou désire-t-on aussi leur fournir un conseiller automatisé qui les guidera (par exemple ATCoach)? Si on choisit un système incorporant un conseiller automatisé, il faudra évaluer si la formation dispensée par le conseiller satisfait aux besoins des élèves contrôleurs dans le contexte canadien.

Ensuite, il est souhaitable de déterminer dans quelle situation le système sera exploité. Désire-t-on des postes isolés ou des postes reliés par réseau avec transfert d'un secteur à un autre?

Lorsque ces questions auront été résolues, il sera possible d'envisager la mise au point d'un système de pseudo-pilote automatisé qui répondra vraiment aux besoins de formation des élèves contrôleurs. Il faudra alors faire un choix pour le développement du pseudo-pilote automatisé : doit-on construire un système qui soit entièrement canadien, ou doit-on utiliser un système réalisé à l'extérieur du Canada? Selon ce que nous avons pu observer, le développement d'un pseudo-pilote automatisé est une tâche de longue haleine. Nous croyons qu'il serait profitable de tirer avantage des développements qui ont été faits ailleurs, peut-être par le biais de transferts de technologie. L'adaptation de ces technologies au contexte canadien recquérerait vraisemblablement un effort important, mais les chances de succès seraient vraisemblablement plus élevées que de développer un système entièrement à partir de zéro. De plus, il pourrait alors être possible de combiner les aspects les plus intéressants de divers systèmes (par exemple un système de langage basé sur des schémas avec un conseiller automatisé).

OUVRAGES CITES

(1) Ratzer, G. F. G., Low cost, high performance ATC simulation available, The ICAO Bulletin, Sept. 1987, p. 26-28

(2) Swan, P., ATS Inc., Communication personnelle, Montréal, 20 juillet 1990

(3) Long, G., Federal Aviation Administration, Communication personnelle, dans le cadre de la conférence Speech Tech '90, organisée par Media Dimensions, Inc., 24-26 avril 1990, New-York, NY

(4) Long, G., Federal Aviation Administration, Communication personnelle, Montréal, QC, juillet 1990

(5) Grady, M. W., Porter, J. E., Satzer, W. J., Speech understanding in air intercept controller training system design, Technical Report NAVTRAEQUIPCEN 78-C-0044-1, Naval Training Equipment Center, Orlando, FLA, January 1979, 68 p.

(6) Anders, R. M., Grady, M. W., Nowell, L. H., A laboratory system for air intercept controller training, Technical Report 78-C-0053-1, Naval Training Equipment Center, Orlando, FLA, January 1979, 71 p.

(7) Klein, P., Making the skies friendlier, Sun Tech Journal, May/June 1990, p. 117-118

(8) Koble, H., ITT, Communication personnelle, dans le cadre de la conférence Speech Tech 90, organisée par Media Dimensions, Inc., 24-26 avril 1990, New-York, NY

(9) Meyn, C., Logicon, Communication personnelle, 10 juillet 1990, San Diego, CA

(10) Scardina, J. A., Ryberg, P., Gerstenfeld, A., Future ATC automation aids based upon AI technology, Proceedings of the IEEE, Special Issue on Air Traffic Control, vol. 77, no. 11, Nov. 1989, p. 1625-1633

(11) Gerstenfeld, A., Speech Recognition integrated with ATC simulation, ICAO Bulletin, May 1988, p. 22-23

(12) Gerstenfeld, A., Air traffic control training: in full console or PC, The Society for Computer Simulation, Simulators Series, Edited by A. Sharon and M. R. Fakory, vol. 22, no. 2, p. 316-318

(13) Olorenshaw, L., Speech Systems Inc., Communication personnelle, dans le cadre de la conférence Speech Tech '90, organisée par Media Dimensions, Inc., 24-26 avril 1990, New-York, NY

(14) Benson, P., Vensko, G., A spoken language understanding system for phraseology training of air traffic controllers, Speech Technology, Media Dimensions, vol. 5, no. 1, Oct/Nov 1989, p. 64-69

(15) Falzon, P., Ergonomie cognitive du dialogue, Presses Universitaires de Grenoble, Grenoble, France, 1989, 175 p.

(16) Néel, F., Matrouf, K., Gauvain, J.L., Mariani, J., Reconnaissance vocale et application en aéronautique, Notes de Recherches, 57e Congrès de l'ACFAS, Montréal, 1989, p.37-42

(17) Matrouf, K., Néel, F., Gauvain, J.L., Mariani, F., Adaptative syntax representation in an oral task-oriented dialogue for air-traffic controller training, Proceedings of Eurospeech 89, Paris, Sept. 1989, vol. I, p. 187-190

(18) Matrouf, K., Néel, F., Mariani, J., Système de dialogue orienté par la tâche: une application en aéronautique, Journal d'Acoustique, Tome 2, no. 1, mars 1989, p. 85-93

(19) Matrouf, K., Gauvain, J. L., Néel, F., Mariani, J., Adapting probability-transitions in DP matching process for an oral task-oriented dialogue, Proceedings of the 1990 IEEE International Conference on Acoustics, Speech and Signal Processing, Albuquerque, New Mexico, USA, April 3-6, 1990, p.569-572

(20) Asbury, M.J.A., The application of DVI within an air traffic system, The Official Proceedings of Speech Tech 88, Media Dimensions, Inc., New-York, NY, April 26-28, p. 242-246

(21) Harrison, J. A., Should speech input/output technology be applied to ATC simulators and operational systems, ICAO Bulletin, May 1984, p. 29-33

(22) Starr, A., Frankish, C., Human factors aspects of direct voice input in air traffic control operations, CAA contract no. 7D/S/423/6, copy no. 36, 1986, 81 p.

(23) Transport Canada, Air Traffic Services, Air Traffic Control Manual of Operations, à jour au 24 août 1989

(24) Institut de Formation de Transports Canada (Cornwall), Midtown ACC, procedures and agreements, DBE-0584, SN 2-2-03-01, June 1989, 31 p.

(25) Plutchak, B., Voice operated status boards for the CATCC, The Official Proceedings of Speech Tech '88, Media Dimensions, Inc., April 26-28, New-York, NY, p. 71-73

(26) Minsky, M., A framework for representing knowledge, tiré de Readings in Knowledge representation, R. J. Brachman et H. J. Levesque (Ed.), Morgan Kaufman Publishers, Inc., Los Altos, Ca, 1985, p. 245-262

(27) Abelson, R., Schank, R., Scripts, Plans, Goals and Understanding, Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1977, p. 41

(28) Code d'épellation en radio-téléphonie,, dans Télécommunications aéronautiques, Volume 2, Annexe 10 à la Convention relative à l'Aviation Civile Internationale, Avril 1985, 4e édition

(29) Fiset, J. Y., Descout, R., Robert, J. M., Optimisation de modèles de langage basés sur des schémas, Actes des XVIIIèmes Journées d'Etudes sur la Parole, Montréal, 28-31 mai 1990, p. 237-241

(30) Holden, J. M., Two helicopter flight tests of voice command, Official Proceedings of Military and Government Speech Tech '89, Media Dimensions Inc., Arlington, VA, Nov. 13-15, 1989, p. 31-35

(31) Pallett, D. S., Performance assessment of automatic speech recognizers, Journal of Research of the National Bureau of Standards, Vol. 90, No. 5, September-October 1985, p. 371-387

(32) Poock, G. K., Speech recognition research, application and international efforts, Proceedings of the Human Factors Society, 30th Annual Meeting, 1986, p. 1278-1283

(33) Baker, J. D., Pinto, D. F., Optimal and suboptimal training strategies for automatic speech recognition in noise, and the effects of adaptation on performance, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Tokyo, 1986, p. 745-748

(34) Rosenhoover, A., Communication personnelle, présentée dans le cadre de la conférence Speech Tech '90, 24-26 avril 1990, New-York, NY

(35) Luzzati, D., Analyse linguistique d'un corpus de dialogues hommes-machines, Centre de Recherche en Morphosyntaxe du Français Contemporain, Université de Paris III - Sorbonne Nouvelle, Tome 1, 1988, p. 95-97

(36) Calliope, Synthèse de la parole, chap. XV de La parole et son traitement automatique, J. P. Tubach (Ed.), Collection Technique et Scientifique des Télécommunications, Masson, 1989, p. 408-454

(37) Muller, N.J., Voice compression basics: How ADPCM increases channel capacity, Journal of Data & Computer Communications, vol. 2, no. 4, Spring 1990, p.32-38

(38) Olive, J. P., Nakatani, L. H., Rule-synthesis of speech by word concatenation: A first step, Journal of the Acoustical Society of America, vol. 55, No. 3, March 1974

(39) Young, S. J., Synthesis by rule of prosodic features in Word Concatenation Synthesis, International Journal of Man-Machine Studies, vol. 12, 1980, p. 241-258

(40) Eady, S. J., Dickson, B. C., Urbanczyk, S. C., Claywards, J. A. W., Wynrib, A. G., Pitch assignment rules for speech synthesis by word concatenation, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3, 1987, p. 1473-1476

(41) Klatt, D. H., Review of text-to-speech conversion for English, Journal of the Acoustical Society of America, vol. 82, No. 3, Sept. 1987, p. 737-793

(42) Calliope, Evaluation de la qualité de la parole codée et synthétique, chap. XVI de La parole et son traitement automatique, J.P. Tubach (Ed.), Collection Technique et Scientifique des Télécommunications, Masson, Paris, 1989, p. 455-488

(43) Streeter, L. A., Applying speech synthesis to user interfaces, tiré de Handbook of Human-Computer Interaction, M. Helander (Ed.), Elsevier Science Publishers B.V., 1988. p.321-343

(44) Schwab, E. C., Nusbaum, H. C., Pisoni, D. B., Some effects of training on the perception of synthetic speech, Human Factors, vol. 27, no. 4, 1985, p. 395-408

(45) Simpson, C. A., McCauley, M. E., Roland, E. F., Ruth, J. C., Williges, B. H., System design for speech recognition and generation, Human Factors, 1985, vol. 27, no. 2, p. 115-141

(46) Simpson, C. A, Occupational experience with a specific phraseology: Group differences in intelligibility for synthesized and human speech, Journal of the Acoustic Society of America, vol. 58, Supplement no. 1, 1975, DD7

(47) Logan, J.S., Pisoni, D.B., Preference judgements comparing different synthetic voices, dans Research on Speech Perception, Progress Report No. 12, Speech Research Laboratory, Indiana University, Bloomington, IN, 1986, p. 262-289

(48) Voiers, W.-D., Diagnostic acceptability measure for speech communication systems, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Hartford, Connecticut, USA, May 9-11, 1977, p. 204-207

(49) Pisoni, D. B., Greene, B. G., Logan, J. S., An overview of ten years of research on the perception of synthetic speech, Proceedings of ESCA Tutorial Day and Workshop on Speech Input/Output Assessment and Speech Databases, Noordwijkerhout, The Netherlands, 20-23 sept. 1989, p. 1.1.1-1.1.4

(50) Luce, P. A., Feustel, T. C., Pisoni, D. B., Capacity demands in short-term memory for synthetic and natural speech, Human Factors, vol. 25, no. 1, 1983, p. 17-32

(51) Waterworth, J. A., The cognitive effort of processing synthetic speech, Eleventh International Symposium on Human Factors in Telecommunications, Sept. 9-13, 1985, CCETT, 35510 Cesson Sevigne, France, A1, p. 1-8

(52) Moody, T. S., Joost, M. G., Synthesized speech, digitized speech, and recorded speech: a comparison of listener comprehension rates, Proceedings of the American Voice I/O Society, 1986, p. 263-280

(53) Nusbaum, H. C., Pisoni, D. B., Human factors issues for the next generation of speech recognition systems, The Official Proceedings of Speech Tech '86, New-York, NY, p. 140-144

(54) Nusbaum, H. C., Pisoni, D. B., Human factors issues for the next generation of speech recognition systems, dans Research on Speech Perception Progress Report No. 12, Speech Research Laboratory, Indiana University, Bloomington, IN, 1986, p. 403-412